MEDDELANDEN

FRÅN

LUNDS ASTRONOMISKA OBSERVATORIUM

SERIE II

N:r 23

ÜBER DIE STERBLICHKEIT IN SCHWEDEN 1886—1914

EINE STABILITÄTS- UND KORRELATIONSUNTERSUCHUNG

VON

FREDRIK ESSCHER

VERLAG SCIENTIA

LUND 1920

ÜBER
DIE STERBLICHKEIT IN SCHWEDEN
1886—1914

EINE STABILITÄTS- UND KORRELATIONSUNTERSUCHUNG

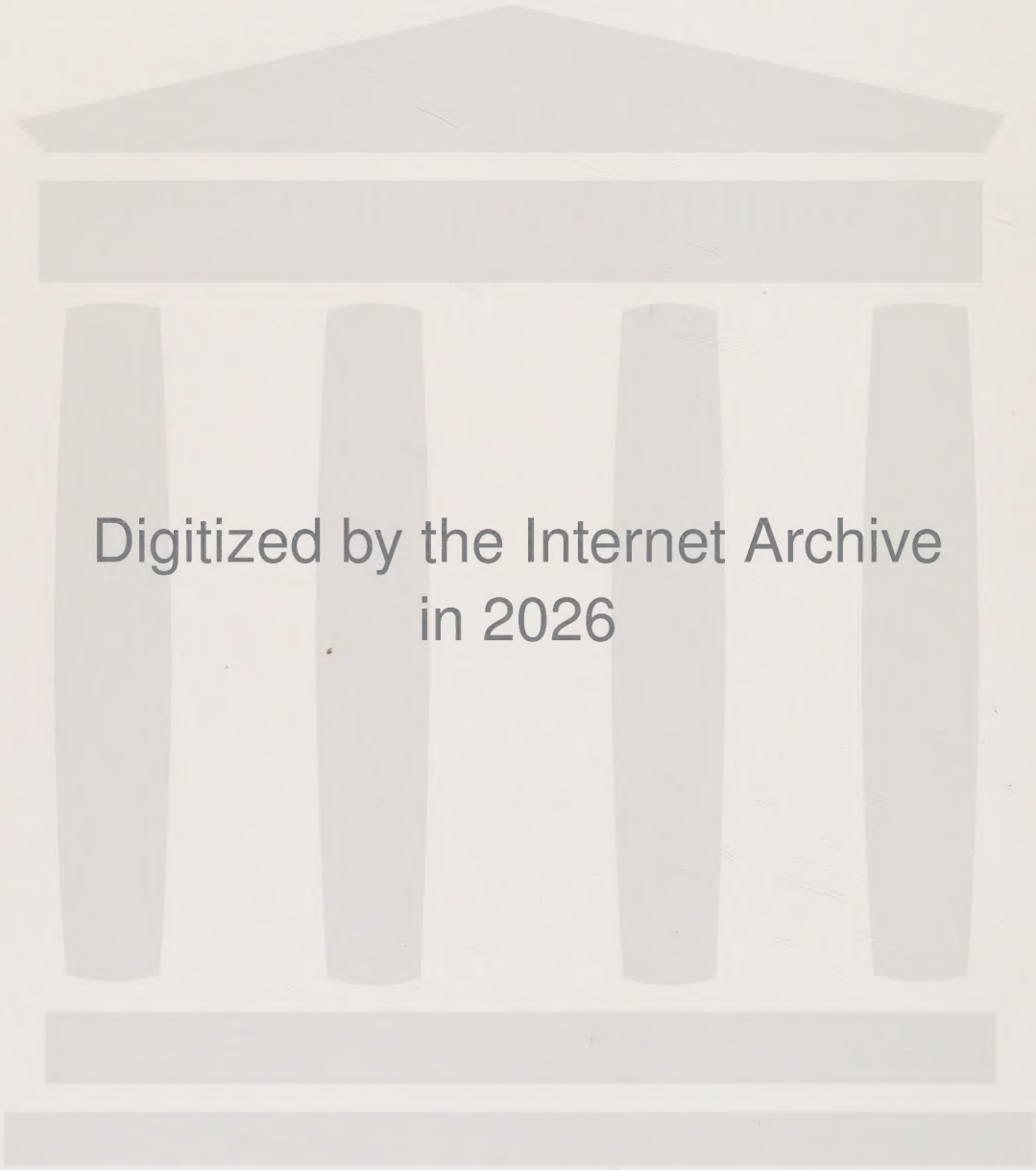
VON

R. UNIV.
BIBLIOTHEEK
LEIDEN

FREDRIK ESSCHER
LIC. PHIL. Yst.

WIRD ZUR ERLANGUNG DER DOKTORWÜRDE MIT GENEHMIGUNG DER PHILOSOPHISCHEN FAKULTÄT ZU LUND ZUR ÖFFENTLICHEN BEURTEILUNG AM 4. DEZEMBER UM 10 UHR VORM. IM HÖRSAAL VI VORGELEGT

VERLAG SCIENTIA
LUND 1920


Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41553402_0046

MEDDELANDEN
FRÅN LUNDS ASTRONOMISKA OBSERVATORIUM
SERIE II. N:r 23

ÜBER
DIE STERBLICHKEIT IN SCHWEDEN
1886—1914

EINE STABILITÄTS- UND KORRELATIONSUNTERSUCHUNG

R. UNIV.
BIBLIOTHEEK
LEIDEN

VON

FREDRIK ESSCHER

VERLAG SCIENTIA
LUND 1920

DRUCK VON LÜTCKE & WULFF, HAMBURG

Einleitung.

Über die Stabilität der Sterblichkeit in den verschiedenen Altersstufen liegen gegenwärtig nur wenige Arbeiten vor. Größere Stabilitätsuntersuchungen sind, soviel ich weiß, nur von PEEK[1]), BOHLMANN[2]) und BLASCHKE[3]) vorgenommen worden. Die beiden letzten, deren Untersuchungen auf versicherungsstatistischem Material gegründet sind, haben in den meisten Altersstufen fast maximale Stabilität gefunden, und zu ungefähr denselben Ergebnissen ist PEEK bei der Untersuchung der Sterblichkeit der männlichen Bevölkerung in den Niederlanden während des Jahrzehnts 1880—1889 gekommen. Hieraus könnte es scheinen, als ob das Problem gelöst sei, und wahrscheinlich hat gerade dieser Umstand dazu beigetragen, daß ähnliche Untersuchungen mit neuem Material nicht vorgenommen worden sind. Die Frage kann jedoch, wegen ihres biologischen Interesses und ihrer großen Bedeutung für das Versicherungswesen, besonders für die Theorie des Risiko, einer erneuten Behandlung wert sein.

Die folgende Untersuchung basiert auf den offiziellen Angaben über die Sterblichkeit in Schweden während der Zeit von 1886 bis 1914. Obwohl die von den Sterblichkeitsverhältnissen der allgemeinen Bevölkerung erhaltenen Resultate natürlich nicht direkt innerhalb des Versicherungswesens anzuwenden sind, so können doch Schlüsse gezogen werden, die auch für ein so selektives Material, wie ein Versicherungsbestand es darstellt, Geltung besitzen. Von biologischem Standpunkt aus gesehen, liefert eine Untersuchung der Sterblichkeitsverhältnisse eines ganzen Volkes die gemeingültigsten Ergebnisse.

Im folgenden soll teils eine Stabilitäts-, teils eine Korrelationsuntersuchung durchgeführt werden. Erstere bekräftigt in gewisser Hinsicht die Ergebnisse früherer Untersuchungen, insbesondere die von PEEK, dessen Arbeit hinsichtlich des Materials

---

[1]) Das Problem des Risiko in der Lebensversicherung. Zeitschr. f. Versicherungsrecht u. Wissenschaft, Bd. V, 1899.

[2]) Über angewandte Mathematik usw. Leipzig 1900.

[3]) Die Anwendbarkeit der Wahrscheinlichkeitslehre im Versicherungswesen usw. Wien 1901.

dieser am nächsten steht. Wesentliche Unterschiede machen sich aber auch bemerkbar. Durch die Korrelationsuntersuchung wird beabsichtigt, die korrelativen Eigenschaften der Sterbenswahrscheinlichkeiten kennenzulernen. Die Theorie zeigt außerdem, daß zwischen Stabilität und Korrelation ein enger Zusammenhang besteht, und die Korrelationsuntersuchung kontrolliert und vervollständigt daher die vorangehende Stabilitätsuntersuchung in sehr wirkungsvoller Weise. Ähnliche Korrelationsuntersuchungen sind, nach meinem Wissen, vorher nicht vorgenommen worden.

Von den numerischen Resultaten werden im folgenden nur die direkt für die Diskussion der Stabilität und der Korrelation notwendigen wiedergegeben werden. Wegen der jetzigen hohen Druckerpreise habe ich nämlich davon Abstand nehmen müssen, vollständige Tabellensammlungen beizufügen.

Kap. I.

Die LEXISsche Reihe und die Stabilitätsmaße.

Obwohl in der Fortsetzung die LEXISsche Reihe nicht direkt angewandt werden wird, liegen doch einige von den Ergebnissen, zu welchen sie führt, der folgenden Untersuchung zugrunde. Diese stützt sich außerdem auf dieselben Voraussetzungen, nämlich auf die Annahme einer serienweise variierenden Grundwahrscheinlichkeit, und in diesem Kapitel wird daher zunächst die LEXISsche Reihe kurz besprochen werden. Von LEXIS aufgestellt, ist diese Reihe theoretisch zuerst von V. BORTKIEWICZ[1]) behandelt worden, dann vollständiger von CHARLIER[2]) und — mit besonderer Berücksichtigung der Korrelation zwischen zwei solchen Reihen — von WICKSELL[3]). Später ist die LEXISsche Reihe von BORTKIEWICZ[4]) von neuem einer ausführlichen Analyse unterzogen worden, und gerade einige von seinen Resultaten sollen hier wiedergegeben werden.

Nehmen wir an, daß N Versuchsserien, „sets", jede aus s Versuchen bestehend, ausgeführt sind! Wir lassen p_k [$= 1 - q_k$] die mathematische Wahrscheinlichkeit, die Grundwahrscheinlichkeit, für das Eintreffen des beobachteten Ergebnisses in der k^{ten} Versuchsserie bedeuten, sowie

$$p = \frac{1}{N}\sum\nolimits_k p_k \qquad [p = 1 - q]. \tag{1}$$

Weiter seien p'_k [$= 1 - q'_k$] und p' die p_k und p entsprechenden empirischen Werte.

$$p' = \frac{1}{N}\sum\nolimits_k p'_k \qquad [p' = 1 - q']. \tag{2}$$

Die Dispersion der beobachteten Reihe wird mit σ,

$$\sigma^2 = \frac{1}{N}\sum\nolimits_k (p'_k - p')^2, \tag{3}$$

und das Quadrat der Dispersion der Grundwahrscheinlichkeiten mit η_2 bezeichnet.

$$\eta_2 = \frac{1}{N}\sum\nolimits_k (p_k - p)^2. \tag{4}$$

---

[1]) Das Gesetz der kleinen Zahlen. Leipzig 1898.

[2]) Contributions etc. Meddelande fr. Lunds Astr. Observatorium. Serie I, N:ris 49 u. 59. 1911 u. 1912.

[3]) Some Theorems etc. Svensk Aktuarietidskrift 1916. On the correlation of acting probabilities. Skandinavisk Aktuarietidskrift 1918.

[4]) Homogenität und Stabilität in der Statistik. Skandinavisk Aktuarietidskrift 1918.

Versteht man unter dem Symbol E die mathematische Erwartung, hat man bei genügend großen s-Werten

$$E(\sigma^2) = \frac{N-1}{N}\frac{pq}{s} + \eta_2. \tag{5}$$

Unter derselben Voraussetzung kann bei der praktischen Anwendung $\frac{pq}{s}$ ohne nennenswerten Fehler durch $\frac{p'q'}{s}$ ersetzt und daraus also ein berechneter Wert von η_2, der „wesentlichen Schwankungskomponente", erhalten werden,

$$\eta_2' = \sigma^2 - \frac{N-1}{N} w_2, \tag{6}$$

wo

$$w_2 = \frac{p'q'}{s}. \tag{7}$$

Es wird nun nach der LEXISschen Dispersionstheorie die Stabilität einer Reihe durch Vergleichen der Dispersion σ mit der BERNOULLIschen Dispersion $\sqrt{w_2}$ untersucht. Je nachdem der Quotient

$$Q = \sigma : \sqrt{\frac{N-1}{N} w_2}, \tag{8}$$

der sog. Divergenzkoeffizient, größer, gleich oder kleiner als 1 ausfällt, sagt man, die untersuchte Reihe habe übernormale, normale oder unternormale Dispersion.

Wenn man vom Unterschied zwischen dem berechneten Wert w_2 und dem theoretischen $\frac{pq}{s}$ absieht, erhält man aus Gleichung (5)

$$E(Q^2) = 1 + \eta_2 : \frac{N-1}{N} w_2, \tag{8*}$$

woraus also hervorgeht, daß in der LEXISschen Reihe die mathematische Erwartung von Q^2 stets größer ist als 1. Ist $Q^2 = 1$, oder weicht es nur unbedeutend von 1 ab, so besagt dies, daß keine oder nur unbedeutende Fluktuationen in den Grundwahrscheinlichkeiten vorkommen. Der Fall, wo Q^2 kleiner ist als 1, ist nach diesem Schema ausgeschlossen.

Der Divergenzkoeffizient hängt indessen, wie aus (8*) ersichtlich ist, bei übernormaler Dispersion von der Größe der Beobachtungszahl ab, und er ist daher, wenn man z. B. die Stabilität zweier Reihen vergleichen will, ein sehr unzweckmäßiges Maß. Dieselbe statistische Erscheinung kann verschiedene Q^2-Werte geben, je nachdem die beobachteten Reihen aus „sets" mit verschiedenen großen Beobachtungszahlen bestehen. Ein anderer mit dem Divergenzkoeffizienten verknüpfter Nachteil ist von CHARLIER[1]) nachgewiesen worden und besteht darin, daß der Wert von Q^2 auch von der durchschnittlichen Größe der Grundwahrscheinlichkeiten abhängig ist. Neben dem Divergenzkoeffizienten hat CHARLIER deshalb als Stabilitätsmaß den sog. Störungskoeffizienten eingeführt, definiert durch

$$100\,\varrho = 100\sqrt{\eta_2} : p. \tag{9}$$

[1]) Contributions etc. Meddelande N:r 52.

Dieser gibt somit an, wieviel Prozent die Dispersion der Grundwahrscheinlichkeiten von dem Mittelwert derselben ausmacht. Der Koeffizient ist später von BORTKIEWICZ[1]) unter der Bezeichnung „die relative wesentliche Schwankungskomponente“ aufgenommen worden.

Dieser Koeffizient ist von der Beobachtungszahl vollkommen unabhängig und auch derart von der Größe der Grundwahrscheinlichkeiten, daß, wenn diese sämtlich in derselben Proportion vergrößert oder verkleinert werden, ϱ immer unverändert bleibt. Da die Größen η_2 und p fast nie a priori bekannt sind, müssen sie bei der numerischen Bestimmung von ϱ durch die nach (6) und (2) berechneten Werte η_2' und p' ersetzt werden.

In der folgenden Stabilitätsuntersuchung soll statt des Divergenzkoeffizienten das Quadrat desselben — das praktisch zweckmäßigste und übrigens theoretisch richtigere Maß — und daneben der Störungskoeffizient angewandt werden. Letzterer ist zwar als Maß für die Stärke der störenden Kräfte dem Divergenzkoeffizienten entschieden überlegen, doch muß man bei seiner Anwendung eine gewisse Vorsicht walten lassen. Ohne sich mit weiteren Rechnungen zu befassen, wird man nämlich leicht einsehen, daß der mittlere Fehler von ϱ nicht nur durch die Größe des Koeffizienten und die Länge der Reihe bestimmt sein kann, sondern auch von der Größe der Beobachtungszahl und der Grundwahrscheinlichkeiten abhängen muß. Auch wenn die aus zwei gleich langen Reihen berechneten Störungskoeffizienten gleich groß ausfallen, ist es also nicht sicher, daß ihnen bei der Beurteilung der Stabilität dieselbe Bedeutung zugemessen werden kann. Zwecks effektiver Ausnutzung der Störungskoeffizienten ist es daher stets notwendig, die Sicherheit der Zahlen irgendwie zu schätzen.

Wegen der mathematischen Schwierigkeiten, mit welchen eine exakte Bestimmung von $E(Q^2)$ und des mittleren Fehlers von Q^2 verknüpft ist, schlägt TSCHUPROW[2]) an Stelle des Divergenzkoeffizienten die Größe η_2 als Stabilitätsmaß vor. Da indessen die Approximationen, die bei der Bestimmung des mittleren Fehlers von Q^2 vorgenommen werden, in der Praxis keine Rolle spielen, und da außerdem Q^2 sich bei der numerischen Berechnung mehrerer Formeln als eine äußerst geeignete Hilfsgröße erwiesen hat, bin ich bei dem alten LEXISschen Stabilitätsmaß geblieben.

In der oben zitierten Arbeit von BORTKIEWICZ sind auch die mittleren Fehler von σ^2 und Q^2 berechnet worden.

Mit Hilfe der Formel

$$\varepsilon^2(x) = E(x^2) - E^2(x) \tag{10}$$

bekommt man in der ersten Approximation

$$\varepsilon(\sigma^2) = \frac{\sigma^2}{Q^2}\sqrt{\frac{2}{N-1}[2Q^2-1]} \tag{11}$$

und, wenn w_2 mit $\frac{pq}{s}$ identifiziert wird,

$$\varepsilon(Q^2) = \sqrt{\frac{2}{N-1}[2Q^2-1]}. \tag{12}$$

---

[1]) Homogenität und Stabilität usw.

[2]) Zur Theorie der Stabilität statistischer Reihen. Skandinavisk Aktuarietidskrift 1918—19.

Ohne Schwierigkeit kann auch nachgewiesen werden, daß

$$\varepsilon(p') = \frac{\sigma}{Q\sqrt{N-1}}. \tag{13}$$

Indessen stellt es sich heraus, daß diese Ausdrücke nicht mit den nach der PEARSONschen Theorie der mittleren Fehler berechneten übereinstimmen. Nach dieser hat man die bekannten Formeln

$$\varepsilon(p') = \frac{\sigma}{\sqrt{N}}, \tag{14}$$

$$\varepsilon(\sigma^2) = \sigma^2\sqrt{\frac{2}{N}}, \tag{15}$$

und mit derselben Approximation wie vorher

$$\varepsilon(Q^2) = Q^2\sqrt{\frac{2}{N}} \text{ oder } \varepsilon(Q) = \frac{Q}{\sqrt{2N}}. \tag{16}$$

Berücksichtigt man bei der Berechnung von $\varepsilon(Q)$ auch den mittleren Fehler von w_2, gelangt man zu dem von CHARLIER[1]) abgeleiteten Ausdruck

$$\varepsilon(Q) = \frac{Q}{\sqrt{2N}}\sqrt{1 + \frac{(p-q^2)}{2spq}Q^2}.$$

Diese Abweichungen im Resultat sind sowohl von BORTKIEWICZ als auch von TSCHUPROW beobachtet worden; doch habe ich keinen Versuch zu ihrer Erklärung gefunden. Die Unterschiede haben indes im Grunde genommen ihre Ursache in verschiedenartigen Definitionen der mittleren Fehler, und die Frage kann daher einer näheren Betrachtung wert sein.

Werden die p'_k-Werte in Anbetracht ihrer Größe in Frequenzgruppen geordnet, so setzt die PEARSONsche Theorie voraus, daß die Wahrscheinlichkeit a priori, — sagen wir P_r —, daß ein p'_k-Wert in die r^{te} Frequenzgruppe falle, eine Elementarwahrscheinlichkeit ist oder eine Durchschnittswahrscheinlichkeit im eigentlichen Sinne, so daß, wenn y_r die Anzahl beobachteter p'_k-Werte in der r^{ten} Frequenzgruppe ist und σ_{y_r} die Dispersion von y_r, die mathematische Erwartung von $\sigma^2_{y_r}$ durch

$$E(\sigma^2_{y_r}) = NP_r(1 - P_r)$$

angegeben wird.

Denkt man sich den Versuch wiederholt, so setzt dies voraus, daß die Grundwahrscheinlichkeiten nicht konstant gehalten werden dürfen, sondern daß in jedem Versuch der Zufall die Größe der Grundwahrscheinlichkeiten bestimmen muß.

Bei der Ableitung der mittleren Fehler nach Formel (10) denkt man sich dagegen den Versuch unter genau denselben Versuchsbedingungen wiederholt, so daß also die Grundwahrscheinlichkeiten von Versuch zu Versuch in entsprechenden „sets“ dieselben bleiben. Unter dieser Voraussetzung wird dann stets, wenn die p'_k-Werte in Frequenz-

[1]) Meddelande N:r 52.

gruppen geordnet angenommen werden, die Wahrscheinlichkeit P_r eine konstant zusammengesetzte Durchschnittswahrscheinlichkeit, und die Dispersion von y_r wird unternormal.

Daß die beiden Methoden verschiedene Resultate liefern, kann demnach daraus erklärt werden, daß die eine sich auf die Hypothese von konstanten Versuchsbedingungen gründet — daraus folgt unternormale Dispersion von y_r —, während die andere voraussetzt, daß die Versuchsbedingungen so variieren, daß die Dispersion von y_r normal wird.

Beide Methoden sind offenbar völlig korrekt, aber die mittleren Fehler bekommen verschiedene Bedeutungen. Ein nach der Hypothese von konstanten Versuchsbedingungen bestimmter mittlerer Fehler liefert ein Maß für die einzig und allein durch die zufälligen Fluktuationen in den Differenzen $p'_k - p_k$ verursachte Unsicherheit eines Resultats, während die p_k-Werte selbst als fehlerfrei vorausgesetzt werden. Geht man dagegen von der Hypothese von variierenden Versuchsbedingungen aus, so müssen die p_k-Werte vom Zufall bestimmt werden, und sie bekommen im allgemeinen bei einem neuen Versuch neue Werte. Die auf ein Resultat einwirkenden Fehler sind also teils aus den zufälligen Fehlern der p_k-Werte, teils aus den zufälligen Abweichungen $p'_k - p_k$ zusammengesetzt; der mittlere Fehler ist ein Maß für die durch diese beiden Faktoren gemeinsam verursachte Unsicherheit.

Sind alle p_k-Werte gleich, d. h., sind die Grundwahrscheinlichkeiten in allen N Versuchsserien dieselben, so müssen beide Methoden dieselben Ergebnisse liefern. Diese Annahme besagt, daß in den Formeln (11), (12) und (13) $Q = 1$ gesetzt werden soll. Man findet dann auch eine vollständige Übereinstimmung — bis auf einen Faktor $\sqrt{\frac{N}{N-1}}$ — zwischen diesen Formeln und den nach der PEARSONschen Theorie berechneten, (14), (15) und (16). Das Aufkommen des Faktors $\sqrt{\frac{N}{N-1}}$ läßt sich daraus erklären, daß PEARSON bei der Bestimmung der Dispersion von y_r und des Korrelationsmomentes von y_r und y_s die weniger korrekten Formeln

$$\sigma_{y_r}^2 = y_r \left(1 - \frac{y_r}{N}\right),$$

$$r_{y_r y_s} \sigma_{y_r} \sigma_{y_s} = -\frac{y_r y_s}{N}$$

verwendet hat, anstatt

$$\sigma_{y_r}^2 = \frac{N}{N-1} y_r \left(1 - \frac{y_r}{N}\right),$$

$$r_{y_r y_s} \sigma_{y_r} \sigma_{y_s} = -\frac{N}{N-1} \times \frac{y_r y_s}{N},$$

die die Bedingungen

$$E(\sigma_{y_r}^2) = N P_r (1 - P_r),$$

$$E(r_{y_r y_s} \sigma_{y_r} \sigma_{y_s}) = -N P_r P_s$$

exakt erfüllen.

Kap. II.

Die Stabilitätsmaße nach Elimination der säkularen Variation.

Außer den rein oszillatorischen Fluktuationen der Sterblichkeit kommt während der Zeit 1886—1914 auch ein starkes säkulares Abnehmen vor, und bevor daher eine Stabilitätsuntersuchung durchgeführt werden kann, muß diese sog. säkulare Variation eliminiert werden. Eine Methode dafür, die augenblicklich viel von englischen Statistikern verwendet wird, ist die sog. „variate-difference"-Methode. Diese gründet sich indes auf Voraussetzungen, die, sollte die Methode auf die Sterblichkeit angewandt werden, nicht erfüllt zu sein scheinen. Ich habe es deshalb vorgezogen, die säkulare Variation vermittels einer nach der Methode der kleinsten Quadrate bestimmten Kurve zu eliminieren. Eine solche Kurve bekommt gewissermaßen den Charakter einer Regressionslinie, und ich werde sie daher in der Fortsetzung nach WICKSELL „die säkulare Regressionslinie der Sterblichkeit" nennen.

Die Methode, mit einer Kurve ein säkulares Ab- oder Zunehmen zu eliminieren, ist von CHARLIER[1]) und WICKSELL[2]) angegeben worden, doch vorausgesetzt, daß es linear vor sich gegangen ist. Der Annahme linearer Variation wird indessen im folgenden nicht überall entsprochen, und ich habe daher die CHARLIER-WICKSELLsche Methode auch auf die Fälle ausgedehnt, wo die säkulare Regressionslinie parabolisch von beliebigem Grade angenommen werden kann.

Ganz allgemein besteht das Problem darin, für das Intervall x_1 bis x_N mittels einer Parabel eine Funktion $F(x)$ auszugleichen, die für eine Reihe von Werten, x_1, x_2, $\ldots x_N$, bekannt ist. Gewöhnlich schreibt man dann

$$F(x) = a_0 + a_1 x + a_2 x^2 + \ldots a_{n-1} x^{n-1},$$

und die Konstanten a_r werden nach der Methode der kleinsten Quadrate bestimmt. Stellt sich nun heraus, daß die Funktion nicht in zufriedenstellender Weise durch diese Parabel dargestellt werden kann, sondern daß Glieder höheren Grades mitgenommen werden müssen, so muß eine völlig neue Lösung durchgeführt werden, wobei die alten Konstanten neue Werte erhalten. Eine Lösung nach der Methode der kleinsten Quadrate ist jedoch meistens ziemlich mühevoll, besonders wenn die Zahl der Unbekannten groß ist, und es ist deshalb oft vorteilhaft, $F(x)$ von Anfang an eine solche Form zu verleihen,

---

[1]) Meddelande N:r 52.

[2]) On the correlation of acting probabilities.

daß das Hinzufügen eines neuen Gliedes die Werte der alten Konstanten nicht verändert. Das kann erreicht werden, wenn die Funktion folgendermaßen geschrieben wird[1]):

$$F(x) = b_0 P_0(x) + b_1 P_1(x) + b_2 P_2(x) + \ldots + b_{n-1} P_{n-1}(x), \tag{17}$$

wo $P_r(x)$ ein Polynom r^{ten} Grades ist mit r vorläufig unbestimmten Konstanten $\alpha_{r,r-1}, \ldots \alpha_{r,0}$.

$$\left\{\begin{aligned} P_0(x) &= 1, \\ P_1(x) &= x + \alpha_{10}, \\ P_2(x) &= x^2 + \alpha_{21} x + \alpha_{20}, \\ &\text{— — — — — — — — — — —} \\ P_r(x) &= x^r + \alpha_{r,r-1} x^{r-1} + \alpha_{r,r-2} x^{r-2} + \ldots + \alpha_{r,0}. \end{aligned}\right. \tag{18}$$

Eine Bestimmung der Konstanten $b_0, b_1, \ldots b_{n-1}$ in (17) nach der Methode der kleinsten Quadrate führt zu folgenden Normalgleichungen.

$$\left\{\begin{aligned} [F P_0] &= b_0 [P_0 P_0] + b_1 [P_0 P_1] + \ldots + b_{n-1} [P_0 P_{n-1}], \\ [F P_1] &= b_0 [P_1 P_0] + b_1 [P_1 P_1] + \ldots + b_{n-1} [P_1 P_{n-1}], \\ &\text{— — — — — — — — — — — — — — — —} \\ [F P_{n-1}] &= b_0 [P_{n-1} P_0] + b_1 [P_{n-1} P_1] + \ldots + b_{n-1} [P_{n-1} P_{n-1}]. \end{aligned}\right. \tag{19}$$

Hier bezeichnet [ ] wie gewöhnlich eine Summe über alle x-Werte von x_1 bis x_N. Nehmen wir jetzt einen Augenblick an, daß die n ersten Polynome $P_0, P_1, \ldots P_{n-1}$ bekannt und daß die Konstanten $b_0, b_1, \ldots b_{n-1}$ von (19) her bestimmt sind! In (17) wird ein neues Glied, $b_n P_n(x)$, hinzugefügt, und die α-Koeffizienten in $P_n(x)$ sollen so bestimmt werden, daß bei einer neuen Lösung nach der Methode der kleinsten Quadrate die vorhin berechneten Werte von $b_0, b_1, \ldots b_{n-1}$ unverändert bleiben. Aus dem Gleichungssystem (19) ist ersichtlich, daß diese Bedingung erfüllt wird, wenn das neue Polynom P_n so beschaffen ist, daß

$$\left\{\begin{aligned} [P_0 P_n] &= 0, \\ [P_1 P_n] &= 0, \\ &\text{— — — —} \\ [P_{n-1} P_n] &= 0. \end{aligned}\right. \tag{20}$$

Diese n Gleichungen sind nun auch ausreichend für eine eindeutige Bestimmung der n unbekannten α-Koeffizienten in P_n.

Werden alle Polynome sukzessiv nach diesem Schema bestimmt, so ist nach (20)

$$[P_r P_s] = 0, \text{ wenn } r \neq s \tag{21}$$

[1]) Einen Entwurf zur Lösung ungefähr desselben Problems habe ich einmal, bevor ich noch glaubte ihn anwenden zu können, unter den hinterlassenen Papieren von Cand. phil. S. BERGSTRÖM zu sehen die Gelegenheit gehabt. Später war es mir jedoch unmöglich, wieder in deren Besitz zu gelangen, und ich bin daher außerstande zu entscheiden, inwiefern die Lösungen gemeinsam sein können.

und die Normalgleichungen nehmen folgende einfache Formen an:

$$(22)\quad \left\{\begin{array}{r} [FP_0] = b_0\,[P_0P_0], \\ [FP_1] = b_1\,[P_1P_1], \\ - - - - - - \\ [FP_n] = b_n\,[P_nP_n], \end{array}\right.$$

woraus die Unbekannten unmittelbar erhalten werden können.

Gibt man demnach der Funktion $F(x)$ die Form (17), und bestimmt man die Polynome nach (20), so kann die Ausgleichung sukzessiv vorgenommen werden, und zwar ohne daß die Mitnahme eines neuen Polynoms in (17) auf die vorhergehenden Resultate einwirkt.

Die mittleren Fehler der nach dieser Methode bestimmten Konstanten sind auch äußerst einfach zu berechnen. Der mittlere Fehler einer Beobachtung ist durch

$$(N-n-1)\,\mu^2 = [(F - b_0P_0 - b_1P_1 - \ldots - b_nP_n)^2]$$

definiert, eine Gleichung, die bei Berücksichtigung von (21) und (22) folgende einfache Form annimmt:

$$(23)\qquad (N-n-1)\,\mu^2 = [F^2] - b_0^2\,[P_0P_0] - b_1^2\,[P_1P_1] - \ldots - b_n^2\,[P_nP_n].$$

Die zeitverschwendende Berechnung der Quadratsumme der scheinbaren Fehler bleibt einem folglich hierdurch erspart.

Den mittleren Fehler einer Konstanten b_r erhält man nachher aus

$$(24)\qquad \varepsilon\,(b_r) = \frac{\mu}{\sqrt{[P_rP_r]}}.$$

Es bleibt noch übrig, mit Hilfe von (20) die Polynome, d. h. die α-Konstanten zu bestimmen. Werden die durch (18) definierten Ausdrücke für $P_0, P_1, \ldots P_{n-1}$ eingeführt, so bekommt man, wenn $n > 0$, durch sukzessives Anwenden der Gleichungen in (20)

$$(25)\quad \left\{\begin{array}{r} [P_n] = 0, \\ [xP_n] = 0, \\ - - - - - \\ [x^{n-1}P_n] = 0. \end{array}\right.$$

Führt man ferner den Ausdruck für P_n hier ein und wendet folgende Bezeichnung an:

$$(26)\qquad e_r = \frac{[x^r]}{N},$$

so ergibt sich

$$(27)\quad \left\{\begin{array}{l} e_n + \alpha_{n,n-1}\,e_{n-1} + \alpha_{n,n-2}\,e_{n-2} + \ldots + \alpha_{n,0} = 0, \\ e_{n+1} + \alpha_{n,n-1}\,e_n + \alpha_{n,n-2}\,e_{n-1} + \ldots + \alpha_{n,0}\,e_1 = 0, \\ - \\ e_{2n-1} + \alpha_{n,n-1}\,e_{2n-2} + \alpha_{n,n-2}\,e_{2n-3} + \ldots + \alpha_{n,0}\,e_{n-1} = 0. \end{array}\right.$$

Setzt man weiter

$$\Delta^{(n)} = \begin{vmatrix} e_{n-1},\ e_{n-2},\ \ldots\ 1, \\ e_n,\ e_{n-1},\ \ldots\ e_1, \\ \text{— — — — — —} \\ e_{2n-2},\ e_{2n-3},\ \ldots\ e_{n-1}, \end{vmatrix}$$

so liefert (27) folgende Lösung:

$$\Delta^{(n)} \alpha_{n,n-r} = -\sum_{1}^{n} {}_s\, e_{n+s-1}\, \Delta^{(n)}_{rs}, \tag{28}$$

und damit ist das Problem theoretisch gelöst.

Bei der praktischen Anwendung können indes die Formeln erheblich vereinfacht werden. Infolge von (25) ist

$$[P_n P_n] = [x^n P_n].$$

Wird auf der rechten Seite der Ausdruck für P_n eingeführt und

$$\frac{1}{N}[P_n P_n] = c_n \tag{29}$$

gesetzt, so erhält man

$$c_n = e_{2n} + \alpha_{n,n-1}\, e_{2n-1} + \ldots + \alpha_{n,0}\, e_n. \tag{30}$$

Daraus kann c_n leicht berechnet werden, nachdem die e-Momente und die α-Konstanten bestimmt worden sind.

Da weiter nach (25)

$$[P_n] = 0, \text{ wenn } n > 0,$$

so ist

$$[F P_n] = [(F - m) P_n], \tag{31}$$

wo m eine beliebige Konstante ist. Durch passende Wahl derselben können die zahlenmäßigen Berechnungen oft vereinfacht werden.

Setzt man in (31) den Ausdruck für P_n ein, so erhält man mit Zuhilfenahme der Bezeichnung

$$d_r = \frac{1}{N}[x^r (F - m)], \tag{32}$$

wenn $n > 0$,

$$\frac{1}{N}[F P_n] = d_n + \alpha_{n,n-1}\, d_{n-1} + \ldots + \alpha_{n,0}\, d_0, \tag{33}$$

und wenn $n = 0$,

$$\frac{1}{N}[F P_0] = d_0 + m. \tag{33*}$$

Die Normalgleichungen (22) können nun so geschrieben werden:

$$(34)\quad \begin{cases} c_0 b_0 = d_0 + m, \\ c_1 b_1 = d_1 + \alpha_{10} d_0, \\ c_2 b_2 = d_2 + \alpha_{21} d_1 + \alpha_{20} d_0, \\ \text{— — — — — — — — —} \\ c_n b_n = d_n + \alpha_{n,n-1} d_{n-1} + \ldots \qquad + \alpha_{n,0} d_0, \end{cases}$$

in welcher Form die Lösung für den praktischen Gebrauch mehr geeignet ist. Aus den gegebenen Werten von x und $F(x)$ werden die Momente $e_1, e_2, \ldots e_{2n}$ nach (26) und die Momente $d_0, d_1, \ldots d_n$ nach (32) berechnet. Aus den e-Momenten erhält man die α-Koeffizienten nach (28) und danach die Konstanten $c_0, c_1, \ldots c_n$ aus (30). Die Unbekannten werden sodann aus (34) berechnet.

Diese Lösung kann indessen in einem speziellen Fall noch vereinfacht werden, nämlich, wenn die Funktion $F(x)$ für äquidistante Werte von x gegeben ist. Durch Änderung der Einheit der x-Koordinate und durch Verschiebung des Nullpunktes kann man es dann stets dahin bringen, daß x alle Werte zwischen 1 und N annimmt. Wird danach die Funktion folgendermaßen geschrieben:

$$(35)\qquad F(x) = b_0 + b_1 P_1\left(x - \frac{N+1}{2}\right) + b_2 P_2\left(x - \frac{N+1}{2}\right) + \ldots,$$

so sind offenbar die d-Momente durch

$$(36)\qquad d_r = \frac{1}{N}\sum_1^N x\,(F - m)\left(x - \frac{N+1}{2}\right)^r$$

und die e-Momente durch

$$(37)\qquad \begin{aligned} e_{2r} &= \frac{1}{N}\sum_1^N x\left(x - \frac{N+1}{2}\right)^{2r} \\ e_{2r+1} &= 0 \end{aligned}$$

definiert.

Infolge von (37) bekommt man nun durch Lösung von (27

$$(38)\qquad \begin{array}{llll} \alpha_{10} = 0 & \alpha_{21} = 0 & \alpha_{32} = 0 & \alpha_{43} = 0, \\ & \alpha_{20} = -e_2 & \alpha_{31} = -e_4 : e_2 & \alpha_{42} = -\dfrac{e_6 - e_4 e_2}{e_4 - {e_2}^2}, \\ & & \alpha_{30} = 0 & \alpha_{41} = 0, \\ & & & \alpha_{40} = -\dfrac{{e_4}^2 - e_6 e_2}{e_4 - {e^2}_2} \quad \text{usw.} \end{array}$$

Schon dadurch, daß alle α-Koeffizienten mit ungeraden Indexsummen verschwinden, wird also die numerische Lösung vereinfacht. Man kann aber auch einfache, allgemeine Ausdrücke für die übrigen α-Koeffizienten und die c-Konstanten bestimmen.

Mit Hilfe der Identität

$$\left(x-\frac{N+1}{2}+1\right)^{2r}-\left(x-\frac{N+1}{2}\right)^{2r}=$$

$$=\binom{2r}{1}\left(x-\frac{N+1}{2}\right)^{2r-1}+\binom{2r}{2}\left(x-\frac{N+1}{2}\right)^{2r-2}+\ldots+1$$

erhält man nach Summierung über alle x-Werte von 1 bis N und im Hinblick auf (37)

$$\frac{(N+1)^{2r}-(N-1)^{2r}}{2^{2r}N}=\binom{2r}{2}e_{2r-2}+\binom{2r}{4}e_{2r-4}+\ldots+\binom{2r}{2}e_2+1.$$

Nach dieser Rekursionsformel habe ich die Momente e_2, e_4, e_6 und e_8 sukzessiv bestimmt und nach Einsetzen in (38) gefunden

$$(39)\quad \begin{cases} \alpha_{20}=-\dfrac{N^2-1}{12}, \\ \alpha_{31}=-\dfrac{3N^2-7}{20}, \\ \alpha_{42}=-\dfrac{3N^2-13}{14}, \\ \alpha_{40}=-9\dfrac{N^2-9}{140}\alpha_{20}, \end{cases} \qquad \begin{aligned} &\alpha_{10}=\alpha_{21}=\alpha_{32}= \\ &=\alpha_{30}=\alpha_{43}=\alpha_{41}=0, \end{aligned}$$

und ferner aus (30)

$$(39^*)\quad \begin{cases} c_0=1, \\ c_1=\dfrac{N^2-1}{12}, \\ c_2=\dfrac{N^2-4}{15}c_1, \\ c_3=9\dfrac{N^2-9}{140}c_2, \\ c_4=4\dfrac{N^2-16}{63}c_3. \end{cases}$$

Diese Ausdrücke gelten sowohl für gerade als auch für ungerade N-Werte.

Wenn deshalb die auszugleichende Funktion für alle x-Werte zwischen 1 und N gegeben ist, oder wenn man es durch Änderung der Einheit und Verschiebung des Nullpunktes dahin bringen kann, so wird die numerische Lösung sehr einfach. Aus (39) und (39*) werden die α- und c-Konstanten und aus (36) die d-Momente berechnet, wonach man die Unbekannten b_0, b_1 ... in (35) aus (34) erhält.

Diese Lösung soll jetzt für die Bestimmung der säkularen Regressionslinien angewandt werden.

Man kann sich dabei denken, daß der Funktion $F(x)$ entweder die beobachteten Sterbenswahrscheinlichkeiten $p'_1, \ldots p'_k, \ldots p'_N$ oder die jeweiligen Grundwahrscheinlichkeiten $p_1, \ldots p_k, \ldots p_N$ entsprechen. In beiden Fällen ist die auszugleichende Funktion für alle x-Werte von 1 bis N gegeben, und die Formelsysteme (36), (39) und (39*) können deshalb angewandt werden. Denkt man sich zuerst die säkulare Regressionslinie aus den Grundwahrscheinlichkeiten bestimmt, so erhält man Gleichungen von folgendem Typus:

$$p_k + \xi_k = B_0 + B_1 P_1 \left(k - \frac{N+1}{2}\right) + B_2 P_2 \left(k - \frac{N+1}{2}\right) +. \tag{40}$$

Die Zahl der Glieder in der Gleichung wird so groß angenommen, daß die säkulare Variation vollständig eliminiert wird. ξ_k, die zeitfreie Residuale genannt, stellt dann den wahren Fehler der Grundwahrscheinlichkeit p_k dar.

Sind $\beta_0, \beta_1, \beta_2, \ldots$ die Werte der Unbekannten, die bei einer Lösung nach der Methode der kleinsten Quadrate zu erwarten wären, und ist Ξ_k der scheinbare Fehler, so würde folgende Beziehung herrschen:

$$p_k + \Xi_k = \beta_0 + \beta_1 P_1 + \beta_2 P_2 +. \tag{41}$$

Die Grundwahrscheinlichkeiten p_k sind indessen nicht bekannt, sondern müssen bei der zahlenmäßigen Berechnung durch die beobachteten Sterbenswahrscheinlichkeiten p'_k ersetzt werden. Die dabei erhaltenen Werte der Unbekannten werden mit $\beta'_0, \beta'_1, \beta'_2 \ldots$ und der scheinbare Fehler mit X_k bezeichnet, so daß

$$p'_k + X_k = \beta'_0 + \beta'_1 P_1 + \beta'_2 P_2 +. \tag{42}$$

Die fünf ersten β'-Konstanten erhält man hier nach (34) aus

$$\left\{\begin{aligned} \beta'_0 &= d_0 + m = p', \\ c_1 \beta'_1 &= d_1, \\ c_2 \beta'_2 &= d_2 + \alpha_{20} d_0, \\ c_3 \beta'_3 &= d_3 + \alpha_{31} d_1, \\ c_4 \beta'_4 &= d_4 + \alpha_{42} d_2 + \alpha_{40} d_0. \end{aligned}\right. \tag{43}$$

Die α- und c-Konstanten sind durch (39) und (39*) gegeben. Ebenso das Moment d_r nach (36) durch

$$d_r = \frac{1}{N} \sum\nolimits^k (p'_k - m) \left(k - \frac{N+1}{2}\right)^r. \tag{44}$$

Die Konstante m setzt man am besten dem provisorischen Medium der p'_k-Werte gleich.

Wie äußerst einfach eine numerische Lösung nach dieser Methode ist, wird in Kap. VI durch ein Zahlenbeispiel veranschaulicht.

In der Gleichung (42) zeigt nun der scheinbare Fehler X_k die Abweichung der beobachteten Sterbenswahrscheinlichkeit von der berechneten säkularen Regressionslinie an. Man hat nämlich

$$X_k = p'_k - p' - \beta'_1 P_1 - \beta'_2 P_2 -. \qquad [\beta'_0 = p'].$$

Wird die Dispersion dieser ausgeglichenen Reihe mit σ_A bezeichnet, so daß

$$\sigma^2{}_A = \frac{1}{N} \sum\nolimits^k X^2{}_k, \tag{45}$$

so erhält man laut (21) und (22) nach Quadrierung und Summierung des obengenannten Ausdrucks

$$\sigma^2{}_A = \sigma^2 - c_1 \beta'^2_1 - c_2 \beta'^2_2 - , \tag{46}$$

wo σ, durch (3) definiert, die Dispersion der unausgeglichenen Reihe angibt.

Ebenso ergibt sich aus (41)

$$\sigma^2{}_\Xi = \eta_2 - c_1 \beta^2{}_1 - c_2 \beta^2{}_2 - , \tag{47}$$

wo

$$\sigma^2{}_\Xi = \frac{1}{N} \sum\nolimits^k \Xi^2{}_k,$$

und wo η_2 durch (4) definiert ist.

Man subtrahiert (47) von (46) und erhält:

$$\sigma^2{}_A - \sigma^2{}_\Xi = [\sigma^2 - \eta_2] - c_1 [\beta'^2_1 - \beta^2{}_1] - c_2 [\beta'^2_2 - \beta^2{}_2] - .$$

Die rechte Seite dieser Gleichung besteht aus Differenzen, in welchen die Minuenden durch die beobachteten Sterbenswahrscheinlichkeiten und die Subtrahenden durch die Grundwahrscheinlichkeiten bestimmt sind, so daß die Differenzen selbst Funktionen der zufälligen Abweichungen $p'_k - p_k$ werden. Werden daher die Differenzen durch ihre resp. erwartungsmäßigen Werte ersetzt, so ist zu bilden:

$$E[\sigma^2{}_A - \sigma^2{}_\Xi] = E[\sigma^2 - \eta_2] - c_1 E[\beta'^2_1 - \beta^2{}_1] - c_2 E[\beta'^2_2 - \beta^2{}_2] - . \tag{48}$$

Man beachte, daß die Bildung des Durchschnittswertes nur auf die zufälligen Abweichungen $p'_k - p_k$ Bezug hat, dagegen nicht auf die event. Fehler in den Grundwahrscheinlichkeiten selber!

Nun ist nach (5)

$$E[\sigma^2 - \eta_2] = \frac{N-1}{N} \frac{pq}{s}. \tag{49}$$

Weiter sind nach (22) die β-Konstanten durch

$$c_r \beta_r = \frac{1}{N} \sum\nolimits^k p_k P_r$$

und die β'-Konstanten durch

$$c_r \beta'_r = \frac{1}{N} \sum\nolimits^k p'_k P_r$$

definiert. Man hat also:

$$E(\beta'_r) = \beta_r.$$

Da ferner

$$c_r [\beta'_r - \beta_r] = \frac{1}{N} \sum\nolimits_k (p'_k - p_k) P_r,$$

kann ohne Schwierigkeit nachgewiesen werden, daß

$$c^2_r E[\beta'_r - \beta_r]^2 = c^2_r E[\beta'^2_r - \beta^2_r] = \frac{1}{N^2} \sum\nolimits_k \frac{p_k q_k}{s} P_r P_r.$$

Ohne nennenswerten Fehler kann hier $\frac{p_k q_k}{s}$ durch $\frac{p q}{s}$ ersetzt werden, und da nach (29)

$$\frac{1}{N} \sum\nolimits_k P_r P_r = c_r,$$

so bekommt man:

(50) $$c_r E[\beta'^2_r - \beta^2_r] = \frac{p q}{s N}.$$

Nach Einsetzen von (49) und (50) in (48) ergibt sich daher:

(51) $$E(\sigma^2_A) = \frac{N - u}{N} \frac{p q}{s} + \sigma^2_\Xi,$$

wo u die Zahl der β'-Konstanten, einschließlich β'_0, in der säkularen Regressionslinie angibt.

Wenn $\frac{p q}{s}$ durch w_2 ersetzt und $E[\sigma^2_A]$ mit σ^2_A identifiziert wird, kann also daraus ein indirekt berechneter Wert von σ_Ξ erhalten werden. Es kommt indessen nicht darauf an, diese Dispersion kennenzulernen, sondern die Dispersion σ_ξ der wahren Fehler ξ_k in der Gleichung (40). Nach der Theorie der Methode der kleinsten Quadrate vertritt σ_ξ den mittleren Fehler „einer Beobachtung“ und wird darum aus der Quadratsumme der scheinbaren Fehler nach der bekannten Formel

(52) $$\sigma^2_\xi = \frac{N}{N - u} \sigma^2_\Xi$$

erhalten.

(51) und (52) liefern daher folgende für die praktische Anwendung geltende Beziehung:

(53) $$\frac{N}{N - u} \sigma^2_A = w_2 + \sigma^2_\xi.$$

Diese ist die fundamentale Gleichung, auf welche die ganze Stabilitätsuntersuchung baut. Die Dispersion der ausgeglichenen Reihe setzt sich somit aus den beiden Komponenten w_2 und σ^2_ξ zusammen, von denen die erstere das Quadrat der BERNOULLIschen Dispersion angibt, die letztere ein Maß für die Verbreitung der Grundwahrscheinlichkeiten

um die wahre säkulare Regressionslinie bildet. Je größer diese Verbreitung ist, desto größer wird σ_ξ, während $\sigma_\xi = 0$ ist, wenn alle Grundwahrscheinlichkeiten genau auf der Regressionslinie liegen. In diesem Fall muß $\frac{N}{N-u}\,\sigma^2{}_A$ mit dem Quadrat der BERNOULLIschen Dispersion gleichwertig sein. Wie für die LEXISsche Reihe kann deshalb hier ein Stabilitätsmaß aufgestellt werden, das dem gewöhnlichen Divergenzkoeffizienten vollkommen entspricht und durch

$$Q^2{}_A = \frac{N}{N-u}\,\sigma^2{}_A : w_2 \tag{54}$$

angegeben wird.

Der Störungskoeffizient entspricht dem Ausdruck

$$100\,\varrho_A = 100\,\sigma_\xi : p' \tag{55}$$

und ist also ein relatives Maß für die Verbreitung der Grundwahrscheinlichkeiten um die wahre Regressionslinie.

Die mittleren Fehler der hier abgeleiteten Ausdrücke habe ich nach der von GAUSS[1]) bei der Bestimmung des mittleren Fehlers der „Gewichtseinheit“ verwendeten Methode berechnet. Da diese Methode sich auf dieselben Voraussetzungen stützt, wie die PEARSONsche Theorie, so werden die mittleren Fehler nach der Hypothese von variierenden Versuchsbedingungen bestimmt. Der Sachverhalt erheischt übrigens, daß die Hypothese von konstanten Versuchsbedingungen hier nicht in Betracht kommen darf; denn die Methode der kleinsten Quadrate setzt voraus, daß die Fehler ξ_k rein zufälliger Natur sind, d. h., daß die Größe der Grundwahrscheinlichkeiten vom Zufall abhängig ist.

Wendet man nun die GAUSSsche Methode an, und bezeichnet man mit x_k den der Größe X_k entsprechenden wahren Fehler, so kann mit Leichtigkeit nachgewiesen werden, daß bei der Gestalt, die hier der Gleichung der säkularen Regressionslinie gegeben worden ist, zwischen X_k und x_k folgende Beziehung besteht:

$$X_k = x_k - \frac{P_0}{[P_0\,P_0]}\,[x\,P_0] - \frac{P_1}{[P_1\,P_1]}\,[x\,P_1] - \cdots, \tag{56}$$

und ferner daß

$$[X\,X] = [x\,x] - \frac{[x\,P_0]^2}{[P_0\,P_0]} - \frac{[x\,P_1]^2}{[P_1\,P_1]} - \cdots. \tag{57}$$

Bildet man nun hier den Durchschnittswert, der, zum Unterschied von den vorhin mit dem Symbol E angegebenen und nur für die Abweichungen $p'_k - p_k$ geltenden, mit M bezeichnet wird und sich auf die zufälligen Fehler x_k und X_k bezieht, so erhält man das wohlbekannte Resultat

$$M[X\,X] = (N-u)\,\mu^2, \tag{58}$$

[1]) Theoria combinationis.

wo $\mu^2 = M(x_k^2)$ das Quadrat des „mittleren Fehlers einer Beobachtung“ angibt und infolge von (45) aus

$$\mu^2 = \frac{N}{N-u}\,\sigma^2{}_A \tag{59}$$

zu berechnen ist [vgl. Formel (23)].

Der mittlere Fehler einer Konstanten β'_r wird daher nach (24) und (29) aus

$$\varepsilon(\beta'_r) = \frac{\sigma_A}{\sqrt{c_r\,(N-u)}}, \qquad [r \leq u-1] \tag{60}$$

erhalten.

Nach Quadrierung von (57) und nach erneuter Bildung des Durchschnittswertes ist weiter leicht zu ersehen, daß

$$M[XX]^2 = (2N - 2u + N^2 - 2Nu + u^2)\,\mu^4. \tag{61}$$

Hier ist vorausgesetzt worden, daß die wahren Fehler normal verteilt sind, so daß $M(x^4{}_k) = 3\,\mu^4$ und daß ferner $M(x^2{}_k\,x_l^2) = \mu^4$; $M(x^3{}_k\,x_l) = M(x^2{}_k\,x_l\,x_m) = M(x_k\,x_l\,x_m\,x_n) = 0$ wenn $k \neq l \neq m \neq n$.

Der mittlere Fehler einer Größe y wird jetzt definiert durch

$$\varepsilon^2(y) = M(y^2) - M^2(y),$$

weshalb man den mittleren Fehler von $[XX]$ nach (61) und (58) aus

$$\varepsilon^2[XX] = 2(N-u)\mu^4$$

erhält und demnach gemäß (45) und (59)

$$\varepsilon(\sigma^2{}_A) = \sigma^2{}_A\sqrt{\frac{2}{N-u}}. \tag{62}$$

Bei einigermaßen großen s-Werten kann man wie vorher von der Unsicherheit der Größe w_2 absehen und deshalb den mittleren Fehler von $Q^2{}_A$ aus

$$\varepsilon(Q^2{}_A) = Q^2{}_A\sqrt{\frac{2}{N-u}} \tag{63}$$

berechnen.

Es zeigt sich, daß ein einfacher und leicht zu handhabender Ausdruck für den mittleren Fehler des Störungskoeffizienten ϱ_A schwer abzuleiten ist. Der Störungskoeffizient und der Divergenzkoeffizient stehen jedoch in so enger Beziehung zueinander, daß es meistens nicht nötig ist, den mittleren Fehler von ϱ_A zu kennen, wenn der mittlere Fehler von $Q^2{}_A$ bekannt ist. Bekommt man nämlich keinen signifikativen Wert $Q^2{}_A - 1$, so kann auch nicht ϱ_A einen solchen besitzen. Anderseits muß die Sicherheit in ϱ_A um so größer werden, je höher der Wert des Divergenzkoeffizienten ist. Die Sicherheit von ϱ_A kann daher meistens mit hinreichender Genauigkeit aus dem Verhältnis zwischen dem Divergenzkoeffizienten und seinem mittleren Fehler beurteilt werden.

Obenstehende Resultate sind sämtlich mit Hilfe der Methode der kleinsten Quadrate abgeleitet und gelten daher, solange die Voraussetzungen, auf welchen diese basiert, erfüllt sind oder, mit anderen Worten, solange die Abweichungen ξ_k und Ξ_k, x_k und X_k Fehler rein zufälliger Art sind. Diese Bedingung kann aber im folgenden nicht überall genau erfüllt sein, denn man kann a priori voraussagen, daß die zeitfreien Residualen ihren absoluten Beträgen nach mit den Grundwahrscheinlichkeiten selber abnehmen. Aber auch wenn die theoretischen Voraussetzungen nicht erfüllt sind, so bestimmt dennoch die Methode der kleinsten Quadrate stets die Kurve, der das Beobachtungsmaterial möglichst genau genügt, in der Weise, daß sie die Quadratsumme der Abweichungen, $\sigma^2{}_A$ bzw. $\sigma^2{}_\Xi$, zu einem Minimum macht. Die Gleichungen (46) und (47) und demnach auch (51) bekommen deswegen immer eine reale Bedeutung, während z. B. Gleichung (53) nur unter der Voraussetzung gilt, daß (52) erfüllt ist. Auf Gleichung (51) kann indessen ebensogut wie auf (53) eine Stabilitätsuntersuchung aufgebaut werden, die hinsichtlich des Divergenzkoeffizienten offenbar zu demselben Ergebnis führt. Der Störungskoeffizient sollte in diesem Fall durch

$$100\,\varrho'_A = 100\,\sigma_\Xi : p'$$

angegeben werden und demnach gemäß (52) in folgendem Verhältnis zu dem nach (55) definierten stehen:

$$\varrho'_A = \sqrt{\frac{N-u}{N}}\,\varrho_A.$$

Nun ist u gewöhnlich im Vergleich zu N eine kleine Zahl, und der Unterschied zwischen ϱ'_A und ϱ_A hat daher meistens keine praktische Bedeutung. Beide Stabilitätsmaße, $Q^2{}_A$ und ϱ_A, können daher als gemeingültig aufgefaßt werden, auch wenn die theoretischen Voraussetzungen, auf welche die Methode der kleinsten Quadrate baut, nicht exakt erfüllt sind.

Bei der Bestimmung der mittleren Fehler muß dagegen, welche Methode man auch immer anwenden mag, stets vorausgesetzt werden, daß die Abweichungen rein zufälliger Art sind. Die oben abgeleiteten Formeln für die mittleren Fehler bekommen daher im Gegensatz zu den Stabilitätsmaßen einen hypothetischen Charakter.

Kap. III.

Die Korrelation zwischen zwei von der säkularen Variation befreiten Reihen.

In zwei Reihen, jede aus N Versuchsserien bestehend, sind die Beobachtungszahlen s_1 und s_2, die Grundwahrscheinlichkeiten in den k^{ten} Versuchsserien $p_{1,k}$ bzw. $p_{2,k}$ und die entsprechenden empirischen Werte $p'_{1,k}$ und $p'_{2,k}$. Ferner sind

$$p_1 = \frac{1}{N}\sum\nolimits^{k} p_{1,k}, \qquad p_2 = \frac{1}{N}\sum\nolimits^{k} p_{2,k},$$
$$p'_1 = \frac{1}{N}\sum\nolimits^{k} p'_{1,k}, \qquad p'_2 = \frac{1}{N}\sum\nolimits^{k} p'_{2,k}.$$

Das aus den empirischen Werten bestimmte Korrelationsmoment ist durch

$$\nu_{11} = \frac{1}{N}\sum\nolimits^{k} (p'_{1,k} - p'_1)(p'_{2,k} - p'_2) \tag{64}$$

definiert, und das entsprechende aus den Grundwahrscheinlichkeiten bestimmte Moment durch

$$\eta_{11} = \frac{1}{N}\sum\nolimits^{k} (p_{1,k} - p_1)(p_{2,k} - p_2). \tag{65}$$

Unter der Bedingung, daß eine zufällige Abweichung $p'_{1,k} - p_{1,k}$ in der einen Reihe auf die entsprechende Abweichung $p'_{2,k} - p_{2,k}$ in der anderen keinen Einfluß ausübt oder umgekehrt, kann leicht bewiesen werden, daß

$$E(\nu_{11}) = \eta_{11}. \tag{66}$$

Wir nehmen nun an, daß die Reihen von der säkularen Variation befreit sind, so daß die Abweichungen von den Regressionslinien durch

$$X_{1,k} = p'_{1,k} - p'_1 - \beta'_{1,1} P_1 - \beta'_{1,2} P_2 -,$$
$$X_{2,k} = p'_{2,k} - p'_2 - \beta'_{2,1} P_1 - \beta'_{2,2} P_2 -$$

gegeben sind[1]).

Multipliziert man diese Gleichungen miteinander, und summiert man darauf die Produkte über alle k-Werte, so erhält man hinsichtlich (64), (21) und (22):

$$r_A \sigma_{A_1} \sigma_{A_2} = \nu_{11} - c_1 \beta'_{1,1} \beta'_{2,1} - c_2 \beta'_{1,2} \beta'_{2,2} -, \tag{67}$$

[1]) Hier, wie im folgenden, wird vorausgesetzt, daß beide Reihen in demselben Grade ausgeglichen sind.

wo

$$r_A \sigma_{A_1} \sigma_{A_2} = \frac{1}{N} \sum\nolimits_k X_{1,k} X_{2,k}.$$

σ_{A_1} und σ_{A_2} geben hier die nach (46) berechneten resp. Dispersionen an und r_A den Korrelationskoeffizienten zwischen den beiden ausgeglichenen Reihen. Diesen erhält man also ganz einfach aus (67) nach Berechnung des Korrelationsmomentes ν_{11}.

Denkt man sich die Regressionslinien aus den Grundwahrscheinlichkeiten bestimmt, so sind nach (41)

$$\Xi_{1,k} = p_{1,k} - p_1 - \beta_{1,1} P_1 - \beta_{1,2} P_2 -,$$
$$\Xi_{2,k} = p_{2,k} - p_2 - \beta_{2,1} P_2 - \beta_{2,2} P_2 -.$$

Setzt man

$$r_\Xi \sigma_{\Xi_1} \sigma_{\Xi_2} = \frac{1}{N} \sum\nolimits_k \Xi_{1,k} \Xi_{2,k}, \tag{68}$$

so ergibt sich ebenso wie bei (67) nach Multiplikation und Summierung und bei Berücksichtigung von (65):

$$r_\Xi \sigma_{\Xi_1} \sigma_{\Xi_2} = \eta_{11} - c_1 \beta_{1,1} \beta_{2,1} - c_2 \beta_{1,2} \beta_{2,2} -. \tag{69}$$

Man subtrahiert (69) von (67) und erhält:

$$r_A \sigma_{A_1} \sigma_{A_2} - r_\Xi \sigma_{\Xi_1} \sigma_{\Xi_2} = [\nu_{11} - \eta_{11}] - c_1 [\beta'_{1,1} \beta'_{2,1} - \beta_{1,1} \beta_{2,1}] - \\ - c_2 [\beta'_{1,2} \beta'_{2,2} - \beta_{1,2} \beta_{2,2}] -. \tag{70}$$

Da nun nach (66)

$$E[\nu_{11} - \eta_{11}] = 0,$$

und da ferner nach derselben Methode, die auf S. 17—18 gebraucht worden ist, nachgewiesen werden kann, daß

$$E[\beta'_{1,r} \beta'_{2,r} - \beta_{1,r} \beta_{2,r}] = 0,$$

so bekommt man aus (70):

$$E[r_A \sigma_{A_1} \sigma_{A_2} - r_\Xi \sigma_{\Xi_1} \sigma_{\Xi_2}] = 0$$

oder die für die praktische Anwendung geltende Beziehung

$$r_\Xi \sigma_{\Xi_1} \sigma_{\Xi_2} = r_A \sigma_{A_1} \sigma_{A_2}. \tag{71}$$

Mit Hilfe davon kann nun der Korrelationskoeffizient r_ξ zwischen den zeitfreien Residualen bestimmt werden. Wird der Zusammenhang benutzt, der nach der Gleichung (56) zwischen den wahren und den scheinbaren Fehlern besteht, so sind

$$\Xi_{1,k} = \xi_{1,k} - \frac{P_0}{[P_0 P_0]}[\xi_1 P_0] - \frac{P_1}{[P_1 P_1]}[\xi_1 P_1] - , \tag{72}$$

$$\Xi_{2,k} = \xi_{2,k} - \frac{P_0}{[P_0 P_0]}[\xi_2 P_0] - \frac{P_1}{[P_1 P_1]}[\xi_2 P_1] - . \tag{73}$$

Da eine Lösung der Gleichung (41) nach der Methode der kleinsten Quadrate in sich birgt, daß

$$[\Xi P_r] = 0, \qquad \text{für } r = 0, 1, 2, \ldots u - 1,$$

so erhält man nach Verlängerung von (72) mit $\Xi_{2,k}$ und Summierung:

$$[\Xi_1 \Xi_2] = [\xi_1 \Xi_2].$$

Wird daher (73) mit $\xi_{1,k}$ verlängert und der Ausdruck darauf summiert, so wird

$$[\Xi_1 \Xi_2] = [\xi_1 \xi_2] - \frac{[\xi_1 P_0][\xi_2 P_0]}{[P_0 P_0]} - \frac{[\xi_1 P_1][\xi_2 P_1]}{[P_1 P_1]} - ,$$

woraus man nach Bildung des Durchschnittswertes und in der Annahme, daß

$$M(\xi_{1,k}\, \xi_{2,k}) = r_\xi \sigma_{\xi_1} \sigma_{\xi_2},$$
$$M(\xi_{1,k}\, \xi_{2,l}) = 0, \text{ wenn } k \neq l,$$

folgenden Ausdruck bekommt:

$$M[\Xi_1 \Xi_2] = (N - u) r_\xi \sigma_{\xi_1} \sigma_{\xi_2}$$

oder im Hinblick auf (68)

$$M(r_\Xi \sigma_{\Xi_1} \sigma_{\Xi_2}) = \frac{N - u}{N} r_\xi \sigma_{\xi_1} \sigma_{\xi_2}.$$

Bei der praktischen Anwendung muß man also das Korrelationsmoment der scheinbaren Fehler $\Xi_{1,k}$ und $\Xi_{2,k}$ identifizieren mit dem Korrelationsmoment der zeitfreien Residualen, multipliziert mit dem Faktor $\frac{N-u}{N}$. Verbindet man diese Gleichung mit (71), so erhält man deshalb zur Bestimmung von r_ξ:

$$r_\xi \sigma_{\xi_1} \sigma_{\xi_2} = \frac{N}{N - u} r_A \sigma_{A_1} \sigma_{A_2}. \tag{74}$$

Werden die Divergenzkoeffizienten der beiden Reihen mit Q_{A_1} und Q_{A_2} bezeichnet, und beachtet man, daß nach (53)

$$\sigma^2{}_\xi = \frac{N}{N - u} \sigma^2{}_A - w_2,$$

so kann (74) auch so geschrieben werden:

$$r_A = r_\xi \sqrt{(1 - {}^1/_{Q^2_{A_1}})\,(1 - {}^1/_{Q^2_{A_2}})}. \tag{75}$$

Dies ist der Form nach dasselbe Verhältnis, wie es WICKSELL[1]) zwischen dem direkt berechneten Korrelationskoeffizienten einerseits und dem Korrelationskoeffizienten zwischen den Grundwahrscheinlichkeiten in zwei unausgeglichenen LEXISschen Reihen anderseits gefunden hat.

Zwar kann r_ξ bei einer indirekten Bestimmung nach (74) sehr leicht einen größeren Wert als $+1$ oder einen kleineren als -1 erhalten, der wahre Wert liegt aber stets innerhalb dieser Grenzen, und man bekommt daher, wenn man von der Unsicherheit in den Divergenzkoeffizienten absieht, als theoretischen Maximalwert von $|r_A|$

$$\text{Max.}\,|r_A| = \sqrt{(1 - {}^1/_{Q^2_{A_1}})\,(1 - {}^1/_{Q^2_{A_2}})}. \tag{76}$$

Hieraus und aus (75) geht hervor, daß r_A seinem absoluten Betrage nach stets niedriger ausfällt, als der entsprechende Koeffizient zwischen den zeitfreien Residualen, und daß er nie einen gewissen Maximalwert übersteigen kann. Dieser wird wie der Unterschied zwischen r_A und r_ξ von der Stabilität der Reihen bestimmt. Bei maximaler Stabilität in der einen oder in beiden Reihen kann theoretisch keine Korrelation bestehen, während mit steigenden Werten der Divergenzkoeffizienten der theoretische Maximalwert von r_A wächst und r_A sich r_ξ nähert. Da r_A also von der Stabilität der Reihen abhängig ist, so kann der Koeffizient als Korrelationsmaß völlig irreführend sein, und es muß daher immer nach Möglichkeit der Koeffizient r_ξ berechnet oder, was auf dasselbe hinauskommt, r_A mit seinem Maximalwert verglichen werden. Um aus einer Korrelationsuntersuchung richtige Schlüsse ziehen zu können, ist demnach Kenntnis der Stabilität der Reihen erforderlich. Anderseits gibt aber die Korrelationsuntersuchung selbst in gewisser Hinsicht darüber Auskunft. Kann nämlich Korrelation, positive oder negative, nachgewiesen werden, so setzt dies nach (75) notwendigerweise übernormale Dispersionen in beiden Reihen voraus. Der Grad der Übernormalität kann aber nicht allein durch die Korrelationsuntersuchung bestimmt werden.

Auf Grund dieses Zusammenhanges zwischen Korrelation und Stabilität können indes oft, besonders in zweifelhaften Fällen, von einer Korrelationsuntersuchung wertvolle Angaben erhalten werden. Wenn z. B. die Divergenzkoeffizienten zweier Reihen so klein ausfallen, daß hinsichtlich der mittleren Fehler von der Stabilität nichts Bestimmtes gesagt, eine Korrelation aber, wenn auch eine schwache, nachgewiesen werden kann, so stützt dieses die Annahme, daß die Übernormalität der Dispersionen nicht rein zufälliger Art sein kann, sondern daß die Reihen wirklich störenden Einflüssen ausgesetzt gewesen sind. Um von einer solchen Untersuchung möglichst großen Nutzen zu ziehen, ist es jedoch nötig, den mittleren Fehler des Korrelationskoeffizienten zu kennen.

Setzt man

$$r_A \sigma_{A_1} \sigma_{A_2} = m_A,$$

so ist

$$r_A = \frac{m_A}{\sqrt{\sigma^2_{A_1} \sigma^2_{A_2}}}.$$

[1]) On the correlation of acting probabilities.

Zeigen nun δr_A, δm_A, $\delta \sigma^2_{A_1}$ und $\delta \sigma^2_{A_2}$ die Abweichungen der berechneten Größen von den jeweiligen erwartungsmäßigen Werten an, so erhält man durch Anwendung der gewöhnlichen PEARSONschen Methode zur Bestimmung des mittleren Fehlers eines Korrelationskoeffizienten:

$$(77) \qquad \frac{\delta^2 r_A}{r^2_A} = \frac{\delta^2 m_A}{m^2_A} + \frac{1}{4}\frac{\delta^2 \sigma^2_{A_1}}{\sigma^4_{A_1}} + \frac{1}{4}\frac{\delta^2 \sigma^2_{A_2}}{\sigma^4_{A_2}} - \frac{\delta m_A \delta \sigma^2_{A_1}}{m_A \sigma^2_{A_1}} - \frac{\delta m_A \delta \sigma^2_{A_2}}{m_A \sigma^2_{A_2}} + \frac{1}{2}\frac{\delta \sigma^2_{A_1} \delta \sigma^2_{A_2}}{\sigma^2_{A_1} \sigma^2_{A_2}}.$$

Jetzt ist nach (62)

$$M(\delta^2 \sigma^2_{A_1}) = \varepsilon^2(\sigma^2_{A_1}) = \frac{2\sigma^4_{A_1}}{N-u},$$

$$M(\delta^2 \sigma^2_{A_2}) = \varepsilon^2(\sigma^2_{A_2}) = \frac{2\sigma^4_{A_2}}{N-u}.$$

Weiter kann nach derselben Methode, die bei der Bestimmung dieser Formeln benutzt worden ist, und bei Voraussetzung normaler Korrelation gezeigt werden, daß

$$(78) \quad \begin{cases} M(\delta^2 m_A) = \dfrac{\sigma^2_{A_1}\sigma^2_{A_2} + m^2_A}{N-u}, \\ M(\delta m_A \delta \sigma^2_{A_1}) = \dfrac{2 m_A \sigma^2_{A_1}}{N-u}, \\ M(\delta m_A \delta \sigma^2_{A_2}) = \dfrac{2 m_A \sigma^2_{A_2}}{N-u}, \\ M(\delta \sigma^2_{A_1} \delta \sigma^2_{A_2}) = \dfrac{2 m^2_A}{N-u}. \end{cases}$$

Die Berechnungen sind mit keinerlei direkten Schwierigkeiten verbunden und werden daher wegen ihrer Länge übergangen.

Unter Benutzung dieser Formeln erhält man aus (77) nach Bildung des Durchschnittswertes und nach Vereinfachung

$$(79) \qquad \varepsilon(r_A) = \frac{1 - r^2_A}{\sqrt{N-u}}.$$

Die mittleren Fehler der Koeffizienten der ausgeglichenen Reihen sind also sämtlich — abgesehen davon, daß N durch $N-u$ ersetzt worden ist — ebenso aufgebaut, wie die entsprechenden mittleren Fehler der unausgeglichenen Reihen.

Wie die Ergebnisse des vorhergehenden Kapitels sind diese mit Hilfe der Methode der kleinsten Quadrate abgeleitet, gelten aber teilweise unabhängig von den Voraussetzungen, auf welchen diese beruht. Es sind also sowohl Gleichung (67) als (69) und folglich auch (71) exakt, auch wenn die Abweichungen von den Regressionslinien keinen zufälligen Charakter haben. Aus (71) kann ferner (76) direkt abgeleitet werden, weshalb auch diese Formel als ganz gemeingültig betrachtet werden kann. Von (79) gilt jedoch dasselbe wie von vorher abgeleiteten mittleren Fehlern.

Kap. IV.

Die Stabilitätsuntersuchung.

In den vorhergehenden Kapiteln sind nun sämtliche für die Stabilitäts- und teilweise auch für die Korrelationsuntersuchung notwendigen Formeln abgeleitet worden, und ich gehe deshalb zur numerischen Behandlung des Problems über.

Das Material stammt von „Sveriges Officiella Statistik A". Eine präliminare Untersuchung der allgemeinen Sterbeziffern für Männer und Frauen ergab, daß die Sterblichkeit in den 3—4 letzten Jahrzehnten in ziemlich kontinuierlichem Abnehmen begriffen war, während gleichzeitig die Fluktuationen der Sterbeziffern erheblich verringert worden waren. Dasselbe Verhältnis zeigte sich bei der Untersuchung der Sterblichkeit in einigen willkürlich ausgewählten Altersgruppen. Besonders stark tritt die erhöhte Stabilität nach 1900 hervor, während die Sterblichkeit vor 1900, abgesehen von der rein säkularen Abnahme, keine beträchtlichen Veränderungen aufzuweisen hat. Um diese Verhältnisse näher ins Auge fassen zu können, habe ich für die folgende Untersuchung die Zeitspanne 1886—1914 gewählt, eine um das Jahr 1900 symmetrisch gelegene Periode[1]). Nur aber aus praktischen Gründen, nämlich um die Elimination der säkularen Variation zu erleichtern, habe ich die herausgegriffene Periode eine ungerade Zahl von Jahren umfassen lassen.

Im folgenden ist die Sterblichkeit der Männer und Frauen in verschiedenen Altersstufen, von der 1. bis zur 90., getrennt untersucht worden; weitere Einteilungen habe ich jedoch nicht vorgenommen. Obwohl es sicher interessant wäre, verschiedene Zivilstandesgruppen zu unterscheiden, Kinder, die in und außer der Ehe geboren sind usw., würde indessen eine solche Differenzierung das schon recht umfangreiche Material vervielfachen. Ferner ist niemals eine einzige Altersstufe herausgegriffen worden, sondern immer wenigstens fünf aufeinanderfolgende, und zwar teils um eine sicherere Kenntnis der für die Altersgruppe charakteristischen Sterblichkeitsverhältnisse zu erlangen, teils um die folgende Korrelationsuntersuchung zu erleichtern.

Bezeichnen wir mit

D_x^y die Zahl der im Laufe des Kalenderjahres y im Alter x bis $x+1$ Gestorbenen,
F^y „ „ „ „ „ „ „ „ Geborenen,
L_x^y „ „ „ am Ende „ „ „ Lebenden im Alter x bis $x+1$,

so sind aus den offiziellen Veröffentlichungen für jedes Kalenderjahr von 1886 bis 1914 und für jede einjährige Altersstufe innerhalb folgender Altersgruppen

[1]) Die offiziellen Veröffentlichungen erstreckten sich, als diese Arbeit begonnen wurde, erst bis 1914.

0—10	35—40	65—70
15—20	45—50	75—80
25—30	55—60	85—90

die mit D_x^y, F^y und L_x^y bezeichneten Zahlen für jedes Geschlecht herausgegriffen worden.

Aus den L_x^y-Werten ist die mittlere Bevölkerungszahl nach der Formel

$$M_x^y = \frac{1}{2}(L_x^{y-1} + L_x^y)$$

berechnet worden, wonach die Sterbenswahrscheinlichkeiten für 0-jährige nach[1])

$$q_0^y = \frac{D_0^y}{\frac{3}{4}F^y + \frac{1}{4}F^{y-1}} \tag{80}$$

und für die übrigen Altersstufen gemäß

$$q_x^y = \frac{D_x^y}{M_x^y + \frac{1}{2}D_x^y} \tag{81}$$

bestimmt worden sind.

Für die meisten Altersstufen hätte es sicher genügt, an Stelle der Sterbenswahrscheinlichkeit q_x den zentralen Sterblichkeitskoeffizienten

$$m_x^y = \frac{D_x^y}{M_x^y}$$

zu berechnen. Der Unterschied zwischen m_x und q_x ist in den mittleren Altersgruppen unbedeutend, kann aber in den Kinderjahren und den höheren Altersstufen nicht vernachlässigt werden. Da nun m_x nicht den Charakter einer Wahrscheinlichkeitsgröße besitzt, so würde dessen Anwendung in diesen Altersstufen gewisse Nachteile mit sich führen. Ich habe es deshalb vorgezogen, die Sterbenswahrscheinlichkeiten konsequent anzuwenden.

Die Zeitspanne 1886—1914 ist in zwei Perioden, 1886—1900 und 1901—1914, geteilt, und für jede dieser Teilperioden wie für die ganze Zeitspanne sind aus den nach (80) und (81) berechneten Sterbenswahrscheinlichkeiten, die im folgenden mit p_1', p_2', $\ldots p_N'$ bezeichnet werden, folgende Charakteristiken berechnet worden:

$$p' = \frac{1}{N}\sum\nolimits_k p_k',$$

$$\sigma^2 = \frac{1}{N}\sum\nolimits_k (p_k' - p')^2,$$

$$w_2 = \frac{p'q'}{s^{(0)}},$$

wo

$$1 : s^{(0)} = \frac{1}{N}\sum\nolimits_k \frac{1}{s_k}$$

und wo s_k = dem Nenner im Ausdruck (80) bzw. (81) für das k^{te} Kalenderjahr[2]).

[1]) Siehe: Sveriges Off. Stat. Dödlighets- och livslängdstabeller.

[2]) Betreffs Anwendung des harmonischen Mittels $s^{(0)}$ siehe CHARLIER: „Vorlesungen über die Grundzüge der mathematischen Statistik", Lund 1920, und WICKSELL: „On the correlation of acting probabilities".

Die säkularen Regressionslinien sind für die ganze Zeitspanne parabolisch zweiten Grades, für die beiden Teilperioden linear gedacht, und die β'-Konstanten sind nach (43) und (44) aus

$$c_1 \beta'_1 = \frac{1}{N} \sum\nolimits_k (p'_k - m) \left(k - \frac{N+1}{2}\right),$$

$$c_2 \beta'_2 = \frac{1}{N} \sum\nolimits_k (p'_k - m) \left(k - \frac{N+1}{2}\right)^2 - c_1 d_0$$

bestimmt worden, wo

m = dem provisorischen Medium der p'_k-Werte,

$d_0 = p' - m$,

und wo nach (39*)

$$c_1 = \frac{N^2 - 1}{12},$$

$$c_2 = \frac{N^2 - 4}{15} c_1.$$

Hiernach sind die Charakteristiken der ausgeglichenen Reihen aus

$$\sigma^2{}_A = \sigma^2 - c_1 \beta'^2_1 - c_2 \beta'^2_2 \qquad \text{[Formel (46)]},$$

$$\sigma^2{}_\xi = \frac{N}{N-u} \sigma^2{}_A - w_2 \qquad [\ \text{„}\ (53)],$$

$$Q^2{}_A = \frac{N}{N-u} \sigma^2{}_A : w_2 \qquad [\ \text{„}\ (54)],$$

$$100\, \varrho_A = 100\, \sigma_\xi : p' \qquad [\ \text{„}\ (55)]$$

berechnet worden, wobei für die ganze Zeitspanne $u = 3$, für die beiden Teilperioden $= 2$. Im letzten Falle fällt auch das Glied $c_2 \beta'^2_2$ im Ausdruck für $\sigma^2{}_A$ fort.

Für die Altersgruppe 85—90 konnte indessen nur die Zeitspanne 1891—1914 benutzt werden, denn sowohl die Sterbenswahrscheinlichkeiten der Männer als auch die der Frauen weisen am Ende der 80er Jahre so große Abweichungen von den säkularen Regressionslinien auf (bis zu achtfacher Dispersion), daß sie bei einer Stabilitätsuntersuchung nicht gut mitgenommen werden können. Die Zeitspanne 91—14 ist außerdem in zwei Perioden, 91—00 und 01—14, geteilt worden. In allen übrigen Altersstufen gehen die Abweichungen nirgends über die dreifache Dispersion hinaus, und weitere Ausschlüsse sind daher nicht gemacht worden.

Die Tabelle I liefert für die Perioden 86—14, 86—00 und 01—14 die Mittelwerte der berechneten Sterbenswahrscheinlichkeiten. Diese Ziffern unterscheiden sich nur unbedeutend von den in „Dödlighets- och livslängdstabeller“ für die Jahrzehnte 1891—1900 und 1901—1910 angegebenen, und weitere Kommentare der Tabelle sind daher überflüssig. Sie ist hier nur mitgenommen worden, um die abnehmende Tendenz der Sterblichkeit von der Periode vor bis zur Periode nach 1900 zu veranschaulichen.

Tab. I.

Die Werte von $10^3 p'$.

Alter	Männer			Frauen		
	86–14	86–00	01–14	86–14	86–00	01–14
0— 1	100.67	112.00	88.53	83.21	93.34	72.36
1— 2	26.04	30.49	21.27	24.49	28.92	19.74
2— 3	13.98	17.64	10.05	13.36	16.90	9.56
3— 4	10.42	13.35	7.28	10.05	12.92	6.98
4— 5	8.18	10.54	5.65	8.18	10.54	5.65
5— 6	6.55	8.30	4.69	6.78	8.61	4.82
6— 7	5.53	6.80	4.17	5.47	6.80	4.06
7— 8	4.91	5.94	3.81	4.85	5.99	3.63
8— 9	4.29	5.19	3.33	4.28	5.06	3.45
9—10	3.83	4.52	3.09	3.88	4.59	3.12
15—16	3.29	3.40	3.17	4.19	4.32	4.06
16—17	3.93	3.99	3.87	4.46	4.46	4.45
17—18	4.68	4.82	4.54	4.61	4.54	4.68
18—19	5.19	5.08	5.30	4.89	4.90	4.89
19—20	5.81	5.74	5.88	5.08	5.17	4.99
25—26	6.36	6.58	6.13	5.99	6.02	5.95
26—27	6.41	6.52	6.29	6.04	6.15	5.93
27—28	6.25	6.53	5.95	5.85	5.90	5.80
28—29	6.26	6.49	6.02	5.94	6.06	5.82
29—30	6.40	6.61	6.17	6.09	6.34	5.83
35—36	6.78	7.04	6.50	6.78	7.13	6.41
36—37	6.79	7.16	6.39	6.68	7.09	6.24
37—38	7.20	7.80	6.55	6.77	6.90	6.63
38—39	7.19	7.71	6.64	7.14	7.57	6.67
39—40	7.40	7.80	6.96	7.18	7.71	6.61
45—46	9.37	9.82	8.90	7.89	8.25	7.51
46—47	9.83	10.33	9.29	8.05	8.48	7.58
47—48	10.29	10.76	9.78	8.27	8.33	8.21
48—49	10.60	11.20	9.95	8.58	8.90	8.23
49—50	11.32	11.95	10.64	9.10	9.43	8.75
55—56	15.62	16.21	14.99	12.19	12.63	11.72
56—57	16.69	17.56	15.77	12.94	13.47	12.38
57—58	17.59	18.55	16.57	13.73	14.30	13.12
58—59	18.62	19.41	17.77	14.91	15.57	14.21
59—60	20.27	20.99	19.50	15.71	16.42	14.96
65—66	31.58	33.16	29.89	26.49	27.80	25.10
66—67	33.81	34.94	32.60	28.89	30.64	27.00
67—68	37.72	39.89	35.93	32.03	34.02	29.89
68—69	40.47	41.98	38.85	34.33	35.60	32.98
69—70	43.78	45.70	41.72	38.50	40.89	35.95
75—76	76.85	80.13	73.32	69.07	71.76	66.18
76—77	85.50	90.01	80.66	76.93	79.88	73.78
77—78	93.98	98.37	89.27	84.99	87.43	82.38
78—79	104.42	109.05	99.46	92.28	96.83	87.41
79—80	113.53	117.44	109.35	103.62	107.07	99.92
	91–14	91–00	01–14	91–14	91–00	01–14
85—86	194.52	200.44	190.29	178.22	185.47	173.05
86—87	217.09	225.27	212.24	190.75	195.28	187.50
87—88	226.68	233.36	221.92	212.48	220.17	206.98
88—89	250.80	262.12	242.72	226.64	237.91	218.58
89—90	281.20	304.76	264.38	240.98	245.23	237.94

Tab. II.

Die Werte von $10^2 v$.

Alter	Männer			Frauen		
	86–14	86–00	01–14	86–14	86–00	01–14
0— 1	— 1.54	— 0.54	— 2.52	— 1.64	— 0.48	— 2.67
1— 2	— 2.43	— 1.28	— 4.26	— 2.52	— 1.17	— 4.26
2— 3	— 3.69	— 2.62	— 5.33	— 3.75	— 2.67	— 5.45
3— 4	— 3.97	— 2.84	— 5.76	— 4.04	— 3.20	— 5.42
4— 5	— 3.98	— 2.78	— 5.67	— 3.98	— 2.74	— 5.11
5— 6	— 3.70	— 2.54	— 5.10	— 3.78	— 2.73	— 5.11
6— 7	— 3.19	— 2.38	— 3.84	— 3.43	— 2.20	— 5.50
7— 8	— 2.93	— 2.40	— 3.33	— 3.24	— 2.35	— 3.94
8— 9	— 2.88	— 1.32	— 4.73	— 2.60	— 1.22	— 5.05
9—10	— 2.48	— 1.54	— 3.28	— 2.59	— 1.46	— 4.32
15—16	— 0.53	— 0.20	— 1.30	— 0.41	+ 0.60	— 1.80
16—17	— 0.28	+ 0.18	— 1.36	— 0.19	+ 0.23	— 1.89
17—18	— 0.29	+ 0.51	— 0.50	+ 0.06	+ 0.80	— 1.85
18—19	+ 0.27	+ 0.74	— 0.43	— 0.01	+ 1.09	— 1.27
19—20	+ 0.24	+ 1.25	— 0.44	— 0.18	+ 0.80	— 0.99
25—26	— 0.38	+ 0.32	— 0.53	— 0.15	+ 0.01	— 0.85
26—27	— 0.23	+ 0.71	— 1.29	— 0.26	+ 0.90	— 1.79
27—28	— 0.61	+ 0.39	— 1.76	— 0.32	— 0.81	— 1.00
28—29	— 0.62	— 0.64	— 1.21	— 0.24	+ 0.46	— 0.93
29—30	— 0.31	+ 0.74	— 0.49	— 0.51	— 0.19	— 0.93
35—36	— 0.39	+ 1.00	— 1.20	— 0.62	+ 0.05	— 0.79
36—37	— 0.51	+ 0.45	+ 0.08	— 0.80	— 0.05	— 1.33
37—38	— 1.02	— 0.22	— 0.91	— 0.40	— 0.57	— 1.10
38—39	— 0.76	+ 0.24	— 0.25	— 0.83	— 0.57	— 0.88
39—40	— 0.71	— 0.19	— 0.88	— 1.06	— 0.68	— 1.60
45—46	— 0.75	— 0.35	— 1.77	— 0.73	— 0.91	— 1.07
46—47	— 0.53	+ 0.66	— 0.71	— 0.31	— 0.76	— 1.17
47—48	— 0.60	+ 0.02	— 1.02	— 0.36	— 1.19	— 1.00
48—49	— 0.67	— 0.13	— 0.41	— 0.44	+ 0.13	— 0.53
49—50	— 0.76	— 0.21	— 1.17	— 0.55	— 0.16	— 1.26
55—56	— 0.51	— 0.49	— 0.29	— 0.61	— 0.91	— 0.85
56—57	— 0.59	— 0.04	— 0.25	— 0.63	— 0.88	— 0.64
57—58	— 0.96	— 0.28	— 0.52	— 0.64	— 0.79	— 0.77
58—59	— 0.52	— 0.17	— 0.42	— 0.55	0.00	— 0.89
59—60	— 0.53	— 0.97	— 0.10	— 0.67	— 0.88	— 0.60
65—66	— 0.62	— 0.35	— 0.27	— 0.56	— 0.45	+ 0.34
66—67	— 0.36	— 0.12	+ 0.09	— 0.77	— 0.83	+ 0.01
67—68	— 0.64	— 0.52	— 0.92	— 0.85	— 1.04	— 0.25
68—69	— 0.46	— 0.44	— 0.06	— 0.48	— 0.61	+ 0.02
69—70	— 0.53	— 0.07	— 0.51	— 0.76	— 0.56	— 0.13
75—76	— 0.63	— 0.81	— 0.53	— 0.53	— 0.36	— 0.61
76—77	— 0.64	— 0.44	— 0.14	— 0.48	— 0.27	— 0.26
77—78	— 0.55	— 0.18	— 0.26	— 0.48	— 0.67	— 0.67
78—79	— 0.54	— 0.05	— 0.54	— 0.58	— 0.12	— 0.33
79—80	— 0.37	+ 0.07	— 0.13	— 0.42	— 0.54	+ 0.17
	91–14	91–00	01–14	91–14	91–00	01–14
85—86	— 0.30	+ 0.09	+ 0.07	— 0.44	+ 0.17	— 0.15
86—87	— 0.48	— 0.16	— 0.34	— 0.27	+ 0.09	— 0.14
87—88	— 0.37	+ 0.03	— 0.34	— 0.34	— 0.18	+ 0.24
88—89	— 0.52	+ 0.44	— 0.43	— 0.52	+ 0.25	— 0.11
89—90	— 0.95	— 0.30	— 0.27	— 0.13	+ 0.21	+ 0.23

Tab. III.

Die Werte von Q^2_A und $100\,\varrho_A$.

Alter	Männer						Frauen					
	Q^2_A			$100\rho_A$			Q^2_A			$100\rho_A$		
	86—14	86—00	01—14	86—14	86—00	01—14	86—14	86—00	01—14	86—14	86—00	01—14
0— 1	26.47	25.64	22.30	5.73	5.32	5.61	21.66	18.99	20.11	5.89	5.17	6.10
1— 2	14.09	16.28	8.67	8.88	8.93	7.47	12.76	14.99	7.14	8.83	8.92	7.08
2— 3	15.34	21.92	3.52	12.95	14.06	6.34	9.33	11.78	4.96	10.26	10.47	8.31
3— 4	12.76	16.50	6.05	13.77	14.11	10.64	8.48	10.93	4.19	11.34	11.63	8.81
4— 5	9.97	11.87	6.45	13.51	13.45	12.62	8.87	9.71	7.14	13.06	12.21	13.65
5— 6	5.75	6.97	3.51	11.23	11.35	9.48	6.18	7.26	4.35	11.73	11.63	11.00
6— 7	4.82	5.55	3.69	11.05	11.05	10.47	5.47	5.91	4.87	12.19	11.71	12.95
7— 8	3.26	4.31	2.63	9.08	10.17	5.30	4.85	6.02	2.72	12.08	12.62	9.18
8— 9	2.38	2.37	1.82	7.64	7.03	6.56	4.53	4.28	4.68	12.40	11.16	13.86
9—10	3.19	3.75	2.21	10.22	10.78	8.30	2.40	1.90	3.28	8.25	6.17	11.53
15—16	1.32	1.67	0.91	4.41	6.44	$2.27\,i$	1.63	1.42	1.75	5.51	4.48	5.96
16—17	1.92	1.60	2.23	6.84	5.67	7.72	1.84	1.78	1.12	3.93	6.19	2.16
17—18	2.39	2.26	2.47	7.84	7.58	7.86	2.32	2.55	1.51	7.81	8.81	4.65
18—19	1.66	1.37	2.13	5.18	4.03	6.51	1.45	1.79	1.05	4.52	6.14	1.39
19—20	1.22	1.83	0.83	2.86	5.85	$2.38\,i$	1.42	1.46	0.93	4.33	4.46	$1.71\,i$
25—26	1.64	1.79	1.30	5.17	5.85	3.41	1.30	1.38	1.09	3.62	4.17	1.97
26—27	1.29	1.39	0.93	3.53	4.22	$1.70\,i$	0.90	0.96	0.71	$2.12\,i$	$1.41\,i$	$3.39\,i$
27—28	1.13	1.38	0.81	2.38	4.15	$2.84\,i$	1.21	1.63	0.35	3.06	5.47	$5.30\,i$
28—29	1.23	1.41	0.97	3.23	4.36	$1.06\,i$	1.40	1.42	1.38	4.27	4.46	4.05
29—30	1.76	1.86	1.32	5.88	6.40	3.73	0.90	1.01	0.69	$2.08\,i$	0.74	$3.74\,i$
35—36	1.28	1.10	1.18	3.69	2.19	3.18	1.18	1.03	1.43	2.91	1.22	4.48
36—37	1.21	0.75	1.09	3.24	$3.48\,i$	2.13	1.23	1.38	0.94	3.31	4.70	$1.67\,i$
37—38	1.87	2.01	1.49	6.51	6.88	5.01	0.94	0.48	1.32	$1.79\,i$	$4.92\,i$	3.85
38—39	1.82	1.87	1.22	6.35	6.21	3.34	1.52	1.76	1.36	4.85	5.75	4.14
39—40	1.40	1.57	1.22	4.45	5.28	3.33	0.81	1.25	0.36	$2.99\,i$	3.29	$5.55\,i$
45—46	1.36	1.73	1.02	3.88	5.63	0.98	1.22	1.45	0.95	3.14	4.53	$1.48\,i$
46—47	2.64	2.93	1.33	8.22	8.96	3.70	1.13	1.52	0.65	2.44	4.83	$3.92\,i$
47—48	1.02	1.13	1.01	1.04	2.30	0.56	2.02	2.77	0.47	6.72	9.08	$4.68\,i$
48—49	1.34	1.27	1.25	3.66	3.30	3.10	1.51	1.65	1.32	4.65	5.33	3.68
49—50	1.62	1.43	1.90	4.86	4.07	5.79	1.41	1.72	1.16	4.11	5.52	2.56
55—56	2.40	2.81	2.05	6.44	7.43	5.68	1.66	2.27	0.90	4.74	6.57	$1.83\,i$
56—57	1.91	4.33	1.92	5.17	9.77	5.24	2.07	2.47	1.71	5.91	6.89	4.75
57—58	2.64	4.70	0.71	6.82	10.36	$2.89\,i$	0.96	1.09	0.83	$1.22\,i$	1.64	$2.31\,i$
58—59	1.18	0.78	1.66	2.22	$2.42\,i$	4.28	1.49	1.56	1.36	3.53	4.04	3.28
59—60	2.19	2.06	2.49	5.52	5.22	6.22	1.43	1.94	0.89	3.53	5.15	$1.84\,i$
65—66	2.07	2.76	1.13	4.56	5.72	1.57	2.91	3.70	1.39	6.08	7.12	2.79
66—67	2.75	3.52	1.67	5.74	6.83	3.59	2.93	4.10	0.97	5.95	7.42	$0.69\,i$
67—68	2.52	3.14	1.90	5.12	6.08	3.98	2.37	3.18	1.36	4.84	6.03	2.56
68—69	1.67	1.95	1.26	3.36	4.02	2.08	2.69	3.69	1.53	5.30	6.72	2.93
69—70	0.97	0.85	1.01	$0.77\,i$	$1.58\,i$	0.47	2.55	3.60	0.69	4.91	6.28	$2.19\,i$
75—76	2.13	2.38	1.85	4.01	4.57	3.33	5.03	6.68	2.68	6.89	8.61	4.40
76—77	3.18	4.06	1.44	5.51	6.79	2.35	3.52	4.24	2.50	5.53	6.50	4.05
77—78	2.99	3.63	1.65	5.32	6.45	2.77	3.82	5.60	1.19	5.83	7.84	1.43
78—79	2.16	2.06	2.31	4.08	4.15	3.96	3.95	4.31	2.97	6.04	6.71	4.61
79—80	2.88	2.41	3.31	5.30	4.98	5.30	4.12	5.45	2.40	6.20	7.88	3.80
	91—14	91—00	01—14	91—14	91—00	01—14	91—14	91—00	01—14	91—14	91—00	01—14
85—86	2.27	2.94	1.52	5.35	7.52	3.06	4.70	6.24	3.22	7.80	10.28	5.43
86—87	2.26	2.61	2.01	5.62	7.24	4.51	2.84	4.56	1.26	5.87	9.05	2.02
87—88	1.22	1.31	1.21	2.61	3.48	2.28	2.09	2.23	1.72	4.80	5.52	3.58
88—89	1.00	0.91	0.95	$0.19\,i$	$2.05\,i$	$1.16\,i$	2.11	2.30	1.74	5.25	6.12	3.99
89—90	1.50	1.37	1.40	4.69	4.43	3.81	1.17	1.45	0.82	2.32	4.11	$2.15\,i$

Die Tabelle gibt einem jedoch davon keine Vorstellung, inwiefern die Sterblichkeit innerhalb der jeweiligen Perioden variiert hat, und ich gebe daher in Tab. II den Wert einer Konstanten $10^2 v$, definiert durch

$$10^2 v = 10^2 \beta_1' : p'.$$

Diese gibt also die durchschnittliche jährliche Abnahme bzw. Zunahme der Sterblichkeit an — vorausgesetzt, daß sie linear vorsichgegangen ist —, ausgedrückt in Prozenten des Mittelwerts der Sterbenswahrscheinlichkeiten in der Periode. Man könnte sagen, v sei ein Maß für die relative säkulare Variation.

Wie aus der Tabelle ersichtlich, war diese ungefähr ebenso stark bei Männern und bei Frauen, und am stärksten während der zehn ersten Lebensjahre, mit einem Maximum in den Altersstufen 3—6. Nach den Angaben des Statistischen Zentralbureaus kann dies hauptsächlich der erfolgreichen Bekämpfung der Kinderkrankheiten zugeschrieben werden. Im Gegensatz zur absoluten war die relative Abnahme der Sterbenswahrscheinlichkeiten der 0—1jährigen und der höheren Altersgruppen nicht besonders stark. Ferner geht aus einem Vergleich zwischen den Perioden 86—00 und 01—14 hervor, daß bis an die Gruppe 45—50 die unvergleichbar stärkste Abnahme nach 1900 stattgefunden hat. In den Gruppen 15—20 und 25—30 zeigt die Sterblichkeit sogar zunehmende Tendenz in der ersten Periode. In den übrigen Gruppen, wo die Ziffern sonst ziemlich fluktuierend sind, besteht kein bestimmter Unterschied zwischen der relativen Variation der ersten und der zweiten Periode.

In Tab. III sind die Werte von den Divergenz- und Störungskoeffizienten aufgenommen. Anstatt für jeden $Q^2{}_A$-Wert den mittleren Fehler nach (63) zu berechnen, habe ich ihn für die drei Zeiträume unter der Voraussetzung bestimmt, daß die Dispersionen normal sind, d. h. daß $Q^2{}_A = 1$ ist. Man bekommt dann

$$\varepsilon_0 (Q^2{}_A) = \sqrt{\frac{2}{N-u}}. \tag{82}$$

Dieser mittlere Fehler gibt die Grenzen an, innerhalb welcher die Größe $Q^2{}_A - 1$ mit einer ungefähren Wahrscheinlichkeit von $^2/_3$ zu liegen käme, wenn die Stabilität maximal wäre. Bekommt man daher einen Wert von $Q^2{}_A - 1$, der $\varepsilon_0 (Q^2{}_A)$ um das Dreifache übersteigt, so besteht nur eine geringe Wahrscheinlichkeit dafür, daß die Stabilität der Reihe maximal sein kann, und je größer im Vergleich zum mittleren Fehler $Q^2{}_A - 1$ ist, mit desto größerer Sicherheit kann die Dispersion als übernormal bezeichnet werden. Aus (82) erhält man:

für die Periode 86—14: $\varepsilon_0 (Q^2{}_A) = 0.28$,
„ „ „ 86—00: $= 0.39$,
„ „ „ 01—14: $= 0.41$.

Aus den $Q^2{}_A$-Werten in Tab. III geht nun hervor, daß in den zehn ersten Altersstufen sowohl bei Männern als bei Frauen unzweifelhaft übernormale Dispersionen vorliegen. Dies gilt sowohl für den ganzen Zeitraum 86—14 als auch für die beiden Teilperioden. Dasselbe kann auch mit einigen wenigen Ausnahmen von den Altersstufen

75—80 behauptet werden. In der Gruppe 65—70 und besonders beim weiblichen Geschlecht lassen die Divergenzkoeffizienten auf ziemlich stark übernormale Dispersionen in den meisten Altersstufen in der Periode 86—00 und auch während des ganzen Zeitraums 86—14 schließen, während über die spätere Teilperiode nichts Bestimmtes gesagt werden kann. Auch die ersten Altersstufen in der Gruppe 85—90 zeigen ziemlich hohe Q^2_A-Werte, die jedoch in den folgenden Altersstufen bedeutend kleiner werden. Anlaß dazu geben wahrscheinlich größtenteils die hier schnell abnehmenden „Beobachtungszahlen". Für die übrigen Altersstufen können aus dieser Tabelle kaum bestimmte Schlüsse gezogen werden. Man beachte indessen, daß die Q^2_A-Werte fast durchgehends größer als 1 sind, obwohl die Abweichungen von 1 in der Regel innerhalb der Grenzen der mittleren Fehler fallen.

In einer fünfjährigen Gruppe (abgesehen von den beiden ersten und möglicherweise auch von der letzten) kann jedoch angenommen werden, daß die äußeren Ursachen, die auf die Sterblichkeit Einfluß haben, zum großen Teil für die verschiedenen Altersstufen der Gruppe gemeinsam sind, eine Annahme, deren Richtigkeit durch die folgende Korrelationsuntersuchung bestätigt wird. Da ferner sowohl Beobachtungszahlen als auch die Größe der Sterbenswahrscheinlichkeiten von einer Altersstufe zur andern nicht zusehends variieren, so sind wahrscheinlich die Fluktuationen der Divergenzkoeffizienten wenigstens zum größten Teil rein zufälliger Art. Bildet man daher für eine Gruppe das arithmetische Mittel der Q^2_A-Werte, so kann dies als ein ziemlich zuverlässiges Maß für die Stabilität der Sterbenswahrscheinlichkeiten der betreffenden Altersstufen betrachtet werden. Bezeichnet man das arithmetische Mittel mit $\overline{Q}^2_A$, so hat man also

$$\overline{Q}^2_A = \frac{1}{n}\sum_i Q^2_{A_i},$$

wo n = der Zahl der Altersstufen in der Gruppe ist. Wie bei den Q^2_A-Werten in Tab. III ist der mittlere Fehler von $\overline{Q}^2_A$ unter der Voraussetzung bestimmt worden, daß die Dispersionen normal sind. Da bei maximaler Stabilität keine Korrelation zwischen den verschiedenen Reihen und demnach auch nicht zwischen ihren Dispersionen bestehen kann, so wird der mittlere Fehler bestimmt durch

$$\varepsilon^2_0(\overline{Q}^2_A) = \frac{1}{n^2}\sum_i \varepsilon^2_0(Q^2_{A_i}),$$

woraus man nach (82) erhält:

$$\varepsilon_0(\overline{Q}^2_A) = \sqrt{\frac{2}{n(N-u)}}.$$

In Tab. IV sind die $\overline{Q}^2_A$-Werte und darunter die mittleren Fehler wiedergegeben. Vom fünften Lebensjahre an sind die $\overline{Q}^2_A$-Werte für jede fünfjährige Gruppe, früher aber wegen der hohen Divergenzkoeffizienten in der ersten Altersstufe nur für 1—5jährige, berechnet worden. Die Fluktuationen der Q^2_A-Werte in den Altersgruppen 1—5 und 5—10 sind jedoch so groß, daß die erhaltenen Mittelwerte kaum als für die Stabilitätsverhältnisse in sämtlichen Altersstufen innerhalb der Gruppen charakteristisch empfunden werden können. Da hier aber schon früher übernormale Dispersionen bestimmt nachgewiesen worden waren, sind die Gruppen der Übersichtlichkeit wegen mit eingereiht

worden. Auch in der Altersgruppe 85—90 können die Mittelwerte wegen der abnehmenden Tendenz der Q^2_A-Werte kaum typisch genannt werden. Für die übrigen Gruppen muß die Tabelle dagegen eine ziemlich gute Vorstellung von den Stabilitätsverhältnissen geben.

Tab. IV.

Die Werte von $\overline{Q}^2_A$.

Alter	Männer			Frauen		
	86—14	86—00	01—14	86—14	86—00	01—14
0— 1	26.47	25.64	22.30	21.66	18.99	20.11
1— 5	13.01	16.64	6.17	9.86	11.85	5.86
5—10	3.88	4.59	2.77	4.69	5.07	3.98
15—20	1.70	1.75	1.71	1.63	1.80	1.27
25—30	1.41	1.57	1.07	1.14	1.28	0.84
35—40	1.52	1.46	1.24	1.14	1.18	1.08
45—50	1.60	1.70	1.30	1.46	1.82	0.91
55—60	2.06	2.94	1.77	1.52	1.87	1.14
65—70	2.00	2.44	1.39	2.69	3.65	1.19
75—80	2.67	2.91	2.11	4.09	5.26	2.35
	91—14	91—00	01—14	91—14	91—00	01—14
85—90	1.65	1.83	1.42	2.58	3.36	1.75

$$\varepsilon_0(\overline{Q}^2_A) = \begin{cases} 0.12 \text{ für die Periode } 86—14, \\ 0.18 \text{ „ „ „ } 86—00 \text{ und } 01—14. \end{cases}$$

Hinsichtlich der mittleren Fehler geht hervor, daß die Dispersionen für den ganzen Zeitraum 86—14 sich nur in den weiblichen Altersgruppen 25—30 und 35—40 dem Normalen nähern, während in allen übrigen Gruppen eine bestimmte, wenn auch schwache, Übernormalität vorliegt. Die Tabelle zeigt indes, daß diese Übernormalität in den mittleren Altersstufen der geringen Stabilität in der Periode 86—00 zugeschrieben werden muß. Diese liefert nämlich fast ausnahmslos höhere $\overline{Q}^2_A$-Werte als die ganze Zeitspanne, während entsprechende Werte in der Periode nach 1900 kleiner ausfallen und für das männliche Geschlecht im Alter 25—50 und für das weibliche im Alter 15—70 so klein sind, daß die Stabilität hier fast maximal angenommen werden muß. Obwohl auch in den übrigen Gruppen die $\overline{Q}^2_A$-Werte in der letzten Periode abgenommen haben, liegen doch immer noch mehr oder weniger stark auffallende übernormale Dispersionen vor. Abgesehen von der abnehmenden Tendenz der $\overline{Q}^2_A$-Werte von der ersten bis zur zweiten Periode scheint die Stabilität sonst keinen wesentlichen Veränderungen ausgesetzt gewesen zu sein. In sämtlichen Zeiträumen zeigen nämlich die $\overline{Q}^2_A$-Werte denselben typischen Gang. Der Stabilitätskoeffizient ist am größten in der ersten Altersstufe, fällt darauf ziemlich schnell und erhält in den Gruppen 25—30 und 35—40 seinen Minimalwert, dann steigt er langsam bis zur Gruppe 75—80, wonach er wieder abnimmt. Die Verstärkung der Stabilität, die — aus den $\overline{Q}^2_A$-Werten zu schließen — die letztere Periode zu kennzeichnen scheint, wird im Anschlusse an die ϱ_A-Werte behandelt werden.

Es dürfte von Interesse sein, die hier gewonnenen Resultate mit entsprechenden von PEEK[1]) aus den Sterbeziffern der männlichen niederländischen Bevölkerung für die Zeit von 1880—89 berechneten zu vergleichen. Letztere werden daher in der untenstehenden Tabelle wiedergegeben.

Tab. Va.

Die PEEKschen Werte von Q^2.

Alter	Q^2	Alter	Q^2	Alter	Q^2	Alter	Q^2	Alter	Q^2
0— 1	41.73	15—16	0.55	35—36	1.19	55—56	1.51	75—76	2.34
1— 2	20.88	16—17	0.92	36—37	1.54	56—57	0.79	76—77	0.49
2— 3	13.54	17—18	2.72	37—38	1.59	57—58	1.49	77—78	2.28
3— 4	14.75	18—19	0.90	38—39	1.02	58—59	0.44	78—79	1.88
4— 5	9.99	19—20	1.74	39—40	0.79	59—60	1.14	79—80	1.32
5— 6	6.35	25—26	1.30	45—46	1.14	65—66	2.02	85—86	1.23
6— 7	4.45	26—27	0.83	46—47	0.50	66—67	1.37	86—87	1.56
7— 8	3.76	27—28	2.32	47—48	1.51	67—68	1.25	87—88	2.31
8— 9	1.28	28—29	1.06	48—49	0.86	68—69	0.69	88—89	0.62
9—10	1.42	29—30	2.34	49—50	1.54	69—70	1.21	89—90	0.96

In Tab. Vb sind die $\overline{Q^2}_A$ gegenüberstehenden Mittelwerte berechnet. Da PEEKs Untersuchung einen Zeitraum von zehn Jahren umfaßt, und da keine Ausgleichung der Reihen vorgenommen worden ist, so werden hier die mittleren Fehler bei maximaler Stabilität durch

$$\varepsilon_0\,(\overline{Q^2}) = \sqrt{\frac{2}{5\,(10-1)}} = 0.21$$

angegeben.

Obwohl ein direkter und detaillierter Vergleich zwischen den Resultaten nicht möglich ist, da dabei auf die Größe der Sterbenswahrscheinlichkeiten und vor allem auf die der „Beobachtungszahlen" Rücksicht genommen werden muß, welch letztere von PEEK nicht angegeben worden sind, so sind doch die Resultate im großen und ganzen miteinander vergleichbar. Die Einwohnerzahlen von Schweden und den Niederlanden sind nämlich von fast derselben Höhe, so daß die Verschiedenheit der Resultate nicht in höherem Maße auf Verschiedenheit der „Beobachtungszahlen" zurückzuführen ist.

Tab. Vb.

Die PEEKschen Werte von $\overline{Q^2}$.

Alter	$\overline{Q^2}$	Alter	$\overline{Q^2}$
0— 1	43.73	45—50	1.11
1— 5	11.83	55—60	1.07
5—10	3.45	65—70	1.31
15—20	1.37	75—80	1.66
25—30	1.57	85—90	1.34
35—40	1.23		

Besonders auffallend ist der hohe Wert, den PEEK für den Divergenzkoeffizienten in der ersten Altersstufe erhalten hat. Dieser wird aber wahrscheinlich teilweise durch säkulare Abnahme der Sterblichkeit hervorgerufen. PEEK hat freilich für einige Altersstufen nachgewiesen, daß die Sterblichkeit während der Zeit von 80—89 einigermaßen

[1]) Loc. cit.

konstant geblieben ist; doch gilt dies sicherlich nicht für die 0—1jährigen. Für alle übrigen Altersgruppen findet man dagegen bei einem Vergleich zwischen PEEKs Resultaten und den hier für die Periode 86—00 aus den Sterbenswahrscheinlichkeiten des männlichen Geschlechts berechneten, daß PEEKs Koeffizienten durchgehend kleiner ausfallen, ja, vom 15. Jahre an so klein, daß die Stabilität von PEEK mit Recht für fast maximal gehalten worden ist. Hierin unterscheiden sich also die Resultate voneinander. Während PEEK nur im Alter von 0 bis 10 Jahren übernormale Dispersionen vorfindet, liefert nämlich Tab. IV entschieden übernormale Dispersionen in beinahe allen Gruppen. Dieser Vergleich gilt für die beiden Epochen, die einander in erster Linie entsprechen würden. Die Unterschiede verschwinden aber in gewisser Hinsicht, wenn PEEKs Resultate mit den für die Periode 01—14 berechneten verglichen werden. Danach scheint die Sterblichkeit in den Niederlanden schon in den 80er Jahren eine Stabilität erreicht zu haben, der sich die Sterblichkeit in Schweden erst in den letzten Jahrzehnten genähert hat.

Bisher ist die Stabilität nur auf der Basis der Q^2_A-Werte behandelt worden, während die ϱ_A-Werte außer acht gelassen worden sind. Wenn es nur darauf ankommt zu konstatieren, ob eine Reihe normale Dispersion hat oder nicht, so ist nämlich, wie vorher gesagt, Q^2_A das geeignetste Stabilitätsmaß, während als Maß für die Stärke der störenden Kräfte in Reihen mit übernormalen Dispersionen der Störungskoeffizient vorzuziehen ist. Da aber vom 15. Jahre ab die Übernormalität der Dispersionen ziemlich schwach ausgeprägt ist, so wird die Unsicherheit der Störungskoeffizienten so groß, daß kaum aus den in Tab. III gegebenen Ziffern irgendwelche Schlüsse direkt gezogen werden können. Nur in den Altersstufen 0—10 können die Störungskoeffizienten als ganz zuverlässig betrachtet werden. Hier zeigen indes die Ziffern einen recht interessanten Verlauf. In der ersten Altersstufe beläuft sich die Dispersion der zeitfreien Residualen auf ca. 5 % des Mittelwertes der Sterbenswahrscheinlichkeiten, darauf steigt dieser Wert mit steigendem Alter, erreicht zwischen dem 3. und 5. Jahre ein Maximum, ca. 12 %, und beginnt erst dann langsam abzunehmen. In dieser Hinsicht zeigen die Störungskoeffizienten einen deutlichen Parallelismus zu entsprechenden Zahlen in Tab. II, womit also gesagt ist, daß in den Altersstufen, wo die relative säkulare Abnahme am stärksten gewesen ist, auch die relativ größten Fluktuationen in den zeitfreien Residualen vorgekommen sind. Wahrscheinlich hängt dieses damit zusammen, daß gerade im Alter von 3—5 Jahren die Kinderkrankheiten am schlimmsten wüten. Bemerkenswert ist ferner der Umstand, daß die Stabilität schneller bei Männern als bei Frauen abnimmt, um danach nach dem Minimum schneller zuzunehmen. Sonst besteht kein größerer Unterschied zwischen den Störungskoeffizienten des männlichen und des weiblichen Geschlechts.

Den Gang der Divergenzkoeffizienten und der Störungskoeffizienten in diesen Altersstufen vergleichend zu verfolgen ist sehr interessant. Bei Betrachtung der ersteren wäre man geneigt zu glauben, daß die Stabilität bei den 0—1jährigen am niedrigsten sei, und daß sie darauf mit steigendem Alter ziemlich kontinuierlich zunähme, während in Wirklichkeit der Sachverhalt ein ganz anderer ist. Dieses Beispiel beleuchtet daher in hervorragender Weise die vorhin erwähnten Nachteile, die den Divergenzkoeffizienten als Stabilitätsmaß anhaften. Die schnelle Abnahme der Q^2_A-Werte hängt hier natürlich teilweise von den von Altersstufe zu Altersstufe fallenden „Beobachtungszahlen“ ab, vor allem aber von der schnellen Abnahme der Sterbenswahrscheinlichkeiten.

Obgleich die ϱ_A-Werte in den meisten übrigen Altersstufen ziemlich unsicher bestimmt sind, kann dennoch für jede fünfjährige Gruppe ein Mittelwert gebildet werden, der, ähnlich wie bei $\bar{Q}^2{}_A$, die Stabilitätsverhältnisse der Gruppe charakterisiert. Da aber einige Störungskoeffizienten imaginär ausgefallen sind, so ist hier statt des arithmetischen Mittels ein quadratischer Mittelwert gebraucht worden, definiert durch

$$100\,\bar{\varrho}_A = 100\sqrt{\frac{1}{n}\sum_i \varrho^2{}_{A_i}}.$$

Diese Koeffizienten stehen berechnet in Tab. VI. Es wäre hier natürlich wünschenswert, die mittleren Fehler zu kennen, aber auch wenn eine Formel für $\varepsilon(\bar{\varrho}_A)$ abgeleitet wäre, würde sie für viele Altersgruppen nicht angewandt werden können. Der mittlere Fehler verliert nämlich seine Gültigkeit bei imaginären ϱ_A-Werten, und in allen Altersgruppen, wo solche Werte bei der Bestimmung von $\bar{\varrho}_A$ mit hineinspielen — und gerade hier möchte man am ehesten die Sicherheit der $\bar{\varrho}_A$-Werte wissen —, könnte daher der mittlere Fehler nicht berechnet werden. Dieser kann auch nicht bestimmt werden, wenn die Stabilität als maximal vorausgesetzt wird, denn dieses bedeutet, daß der erwartungsmäßige Wert von ϱ_A und demnach auch von $\bar{\varrho}_A = 0$ ist, und in diesem Fall ist es theoretisch unmöglich, aus $\varepsilon_0(\bar{\varrho}^2{}_A)$, das zuerst berechnet werden muß, $\varepsilon_0(\bar{\varrho}_A)$ zu erhalten. Die Sicherheit der Ziffern kann indessen mit hinreichender Genauigkeit durch einen Vergleich mit entsprechenden $\bar{Q}^2{}_A$-Werten und deren mittleren Fehlern beurteilt werden. Je größer der Koeffizient $\bar{Q}^2{}_A$ ist, um so größer muß die Sicherheit im entsprechenden $\bar{\varrho}_A$-Wert sein. In einer Gruppe mit schwach übernormaler Dispersion wird dagegen der Koeffizient ziemlich unsicher bestimmt. Daher ist es klar, daß den für die mittleren Altersgruppen, besonders für die Periode 01—14, berechneten $\bar{\varrho}_A$-Werten keine zu große Bedeutung zugemessen werden darf.

Tab. VI.

Die Werte von $100\,\bar{\varrho}_A$.

Alter	Männer			Frauen		
	86—14	86—00	01—14	86—14	86—00	01—14
0— 1	5.73	5.32	5.61	5.89	5.17	6.10
1— 5	12.44	13.57	9.60	10.98	10.88	9.79
5—10	9.74	10.20	8.24	11.44	10.90	11.82
15—20	5.70	5.99	5.53	5.40	6.25	3.49
25—30	4.24	5.03	1.64	2.52	3.63	$2.58\,i$
35—40	5.03	4.61	3.52	2.48	2.94	1.92
45—50	4.91	5.49	3.41	4.46	6.09	$1.97\,i$
55—60	5.48	7.49	4.66	4.02	5.22	2.06
65—70	4.26	5.10	2.67	5.44	6.70	1.88
75—80	4.89	5.49	3.69	6.11	7.55	3.83
	91—14	91—00	01—14	91—14	91—00	01—14
85—90	4.22	5.22	3.10	5.50	7.38	3.40

Aus den Ziffern in obenstehender Tabelle geht hervor, daß die Stabilität, mit dem relativen Maß $\bar{\varrho}_A$ gemessen, in beiden Geschlechtern im Alter von 1 bis 10 Jahren am niedrigsten ist. Die Dispersionen der zeitfreien Residualen belaufen sich hier durchschnittlich auf 10—12 % der Mittelwerte der Sterbenswahrscheinlichkeiten. In den übrigen Gruppen findet man sowohl für den ganzen Zeitraum als auch für die Periode 86—00 kleinere, aber sehr konstante $\bar{\varrho}_A$-Werte, die im Durchschnitt 4—6 % betragen. In der Periode nach 1900 zeigen die Störungskoeffizienten bedeutend größere Fluktuationen, auch wenn man von der Gruppe 1—10 absieht. Hier treten in den mittleren Altersgruppen als eine natürliche Folge der vorhin nachgewiesenen fast normalen Dispersionen, besonders beim weiblichen Geschlecht, sehr kleine, ja sogar imaginäre $\bar{\varrho}_A$-Werte auf, während die Störungskoeffizienten in den übrigen Gruppen noch verhältnismäßig starke Fluktuationen der zeitfreien Residualen andeuten. Im großen und ganzen zeigen aber die Störungskoeffizienten viel konstantere Werte als die Divergenzkoeffizienten, was darauf hinzudeuten scheint, daß die störenden Kräfte, wo solche existieren, in allen Altersstufen mit Ausnahme von der 1. bis 10. relativ ebenso stark wirken. Besonders bemerkenswert erscheint es mir, daß die Stabilität bei den 0—1jährigen nicht niedriger ist, als durchschnittlich in den meisten übrigen Gruppen.

Bei einem Vergleich zwischen den Koeffizienten der ersten und der zweiten Periode stellt es sich heraus, daß die Stabilität in den meisten Altersgruppen während der zweiten Periode — in einigen Fällen sogar recht beträchtlich — zugenommen hat. Um diese Verhältnisse näher zu erläutern, ist Tab. VII aufgestellt worden, die den Unterschied zwischen den $100\text{-}\bar{\varrho}_A$-Werten der beiden Perioden für Männer und Frauen liefert. Die Tabelle zeigt, daß die Stabilität bei 0—1jährigen keinen größeren Schwankungen ausgesetzt gewesen ist. Die negativen Zahlen deuten zwar die Möglichkeit verringerter Stabilität an, aber sie sind allzu klein, um eine bestimmte Entscheidung darüber erzwingen zu können. In allen übrigen Gruppen, nur mit einer Ausnahme, hat die Stabilität zugenommen, und zwar besonders stark bei den Frauen. Zum Teil kann die geringe Stabilität in der ersten Periode den bösartigen Influenzaepidemien zugeschrieben werden, die in den 90er Jahren das Land heimsuchten und dann zeitweise die Sterblichkeit sehr stark beeinflußten. Die Verstärkung der Stabilität nach 1900 ist indessen nicht nur ein für die Sterblichkeit charakteristisches Phänomen. Dr. WICKSELL hat mir mündlich mitgeteilt, er habe auf vielen verschiedenen Gebieten der Bevölkerungsstatistik dieselbe Beobachtung gemacht; die Erscheinung scheine für die letzten Jahrzehnte charakteristisch zu sein.

Tab. VII.

Die Werte von
$100\,\bar{\varrho}_A$ (86—00) — $100\,\bar{\varrho}_A$ (01—14).

Alter	Männer	Frauen
0— 1	— 0.29	— 0.93
1— 5	+ 2.97	+ 1.09
5—10	+ 1.96	— 0.92
15—20	+ 0.46	+ 2.76
25—30	+ 3.39	$3.63 - 2.58\,i$
35—40	+ 1.09	+ 1.02
45—50	+ 2.08	$6.09 - 1.97\,i$
55—60	+ 2.83	+ 3.16
65—70	+ 2.43	+ 4.82
75—80	+ 1.80	+ 3.72
85—90	+ 2.12	+ 3.98

Aus den Ziffern in Tab. VI ist ferner untenstehende Tabelle berechnet worden, die für die drei Zeiträume den Unterschied zwischen den $100\text{-}\bar{\varrho}_A$-Werten für Männer und Frauen liefert.

Die kleinen Ziffern in der ersten Altersstufe deuten darauf hin, daß hier zwischen der Stabilität der Sterbenswahrscheinlichkeiten der Männer und Frauen kein größerer Unterschied herrscht. Die positiven Zahlen in der Gruppe 1—5 und die negativen in der Gruppe 5—10 entstehen dadurch, daß die Stabilität — wie schon erwähnt — im Alter von 1—5 Jahren bei Männern schneller abnimmt als bei Frauen, um im Alter 5—10 schneller zuzunehmen. Von der Gruppe 15—20 an nehmen die Ziffern einen ziemlich bestimmten systematischen Gang. In den mittleren Altersstufen, bis zu 55—60 Jahren hinauf, ist die Stabilität bei Frauen am größten, während von der Gruppe 65—70 an das Verhältnis sich umkehrt. Freilich sind diese Ziffern gerade in den mittleren Altersgruppen ziemlich unsicher; es wäre aber sehr unwahrscheinlich, wenn in allen drei Perioden fast sämtliche Differenzen ganz zufällig diesen systematischen Gang aufweisen würden.

Tab. VIII.

Die Werte von $100\,\bar{\varrho}_A$: Männer — $100\,\bar{\varrho}_A$: Frauen.

Alter	86—14	86—00	01—14
0— 1	— 0.16	+ 0.15	— 0.49
1— 5	+ 1.46	+ 2.69	— 0.19
5—10	— 1.70	— 0.70	— 3.58
15—20	+ 0.30	— 0.26	+ 2.04
25—30	+ 1.72	+ 1.40	$1.64 - 2.58\,i$
35—40	+ 2.55	+ 1.67	+ 1.60
45—50	+ 0.45	— 0.60	$3.41 - 1.97\,i$
55—60	+ 1.46	+ 2.27	+ 2.60
65—70	— 1.18	— 1.60	+ 0.79
75—80	— 1.22	— 2.06	— 0.14
	91—14	91—00	01—14
85—90	— 1.28	— 2.16	— 0.30

Die Resultate scheinen aber recht eigentümlich zu sein. In den mittleren Altersstufen, wo die Sterbenswahrscheinlichkeiten der Männer und Frauen sich nur wenig voneinander unterscheiden, wo die Männer aber infolge ihrer Tätigkeit und ihres Berufs in verschiedenartigeren Verhältnissen leben als die Frauen und demnach eine statistisch weniger homogene Masse bilden, müßte man bei dem Antagonismus, der im allgemeinen zwischen Stabilität und Homogenität[1]) vorherrscht, bei den Männern a priori eine höhere Stabilität erwarten als bei den Frauen. Hier stellt sich aber das Gegenteil heraus. Da aber die Stabilität der Sterblichkeit natürlich nicht nur von der Zusammensetzung der Masse bestimmt wird, sondern u. a. auch von physiologischen Faktoren, so wäre es sehr gewagt, sich hier auf Erklärungen einzulassen. Dazu ist übrigens eine eingehendere Differenzierung des Materials nötig, als die hier vorgenommene.

Die in diesem Kapitel behandelten Resultate sind durch das Studium der zweiten Momente und der davon abgeleiteten Koeffizienten gewonnen worden, während höhere Momente außer acht gelassen worden sind. Es wäre aber unzweckmäßig, für so kurze Reihen wie die hier gebrauchten, höhere Momente zu berechnen. Eine eingehendere Untersuchung der Verteilung der Sterbenswahrscheinlichkeiten um die säkularen Regressionslinien kann daher nicht durchgeführt werden. Summarisch könnte zwar die Verteilung mit einfacheren Hilfsmitteln betrachtet werden, aber diese erfordern die Berechnung der Abweichungen der Sterbenswahrscheinlichkeiten von den säkularen Regressionslinien, und das würde wieder eine Arbeit erheischen, die keineswegs die gewonnenen Resultate wettmachen würde. Ich habe deshalb auch von einer Untersuchung dieser Verhältnisse abgesehen und mich mit den Ergebnissen begnügt, zu welchen das Studium der zweiten Momente führt.

[1]) Siehe hierüber BORTKIEWICZ: Homogenität und Stab. usw.

Kap. V.

Die Korrelationsuntersuchung.

Durch die vorhergehende Stabilitätsuntersuchung ist erwiesen, daß die Schwankungen der Sterbenswahrscheinlichkeiten um die säkularen Regressionslinien in den meisten Altersstufen von solchem Charakter waren, daß äußere Faktoren, äußere störende Kräfte die Sterblichkeit beeinflußt haben müssen. Die folgende Korrelationsuntersuchung ist nun in erster Linie vorgenommen worden, um zu ermitteln, wie weit jene Faktoren verschiedenen Altersstufen und Altersgruppen gemeinsam gewesen sind. Eine eingehende Behandlung des Korrelationsproblems würde indessen bei dem vorliegenden umfangreichen Material ungeheure Mühe erfordern, und ich habe mich deshalb auf eine summarische Behandlungsweise beschränken müssen. Dabei habe ich die Absicht gehabt, die Korrelationsverhältnisse zwischen den verschiedenen Altersstufen innerhalb einer Gruppe und das Korrelationsverhältnis zwischen verschiedenen Gruppen mittels einfach berechneter Koeffizienten mit einiger Zuverlässigkeit zu charakterisieren. Die dabei gebrauchten Koeffizienten sind folgendermaßen hergeleitet worden.

Wir nehmen an, daß n Reihen, jede aus N „sets“ bestehend, vorliegen. Weiter sei $p'_{i,k}$ die k^{te} Beobachtung der i^{ten} Reihe und

$$p'_i = \frac{1}{N}\sum\nolimits_k p'_{i,k}.$$

Bildet man nun

$$D_k = \sum\nolimits_i (p'_{ik} - p'_i), \tag{83}$$

so bekommt man hieraus nach Quadrierung

$$D^2_k = \sum\nolimits_i (p'_{i,k} - p'_i)^2 + \sum_{i \neq j}\sum (p'_{i,k} - p'_i)(p'_{j,k} - p'_j)$$

und nach Summierung über alle k-Werte

$$S^2 = \sum\nolimits_i \sigma^2_i + \sum_{i \neq j}\sum r_{ij}\,\sigma_i\,\sigma_j, \tag{84}$$

wo

$$S^2 = \frac{1}{N}\sum\nolimits_k D^2_k \tag{85}$$

und wo

σ_i = der Dispersion der i^{ten} Reihe,

r_{ij} = dem Korrelationskoeffizienten zwischen der i^{ten} und j^{ten} Reihe.

Sind nach Elimination der säkularen Variation σ_{A_i} und $r_{A_{ij}}$ entsprechende Werte für Dispersion und Korrelationskoeffizient und $\beta'_{i,1}$ und $\beta'_{i,2}$ die Konstanten in der Regressionslinie der i^{ten} Reihe, so ist nach (67)

$$r_{ij}\,\sigma_i\,\sigma_j = r_{A_{ij}}\,\sigma_{A_i}\,\sigma_{A_j} + c_1\,\beta'_{i,1}\,\beta'_{j,1} + c_2\,\beta'_{i,2}\,\beta'_{j,2}$$

und demnach

$$\sum_{i \neq j}\sum r_{ij}\,\sigma_i\,\sigma_j = \sum_{i \neq j}\sum r_{A_{ij}}\,\sigma_{A_i}\,\sigma_{A_j} + \tag{86}$$

$$+ c_1\left(\sum\nolimits^{i}\beta'_{i,1}\right)^2 - c_1\sum\nolimits^{i}\beta'^2_{i,1} + c_2\left(\sum\nolimits^{i}\beta'_{i,2}\right)^2 - c_2\sum\nolimits^{i}\beta'^2_{i,2}.$$

Setzt man nun

$$S^2{}_A = S^2 - c_1\left(\sum\nolimits^{i}\beta'_{i,1}\right)^2 - c_2\left(\sum\nolimits^{i}\beta'_{i,2}\right)^2 \tag{87}$$

und achtet man darauf, daß nach (46)

$$\sigma^2{}_{A_i} = \sigma^2{}_i - c_1\,\beta'^2_{i,1} - c_2\,\beta'^2_{i,2},$$

so ergibt sich nach Einsetzung von (86) in (84)

$$\sum_{i \neq j}\sum r_{A_{ij}}\,\sigma_{A_i}\sigma_{A_j} = S^2{}_A - \sum\nolimits^{i}\sigma^2{}_{A_i}. \tag{88}$$

Wird weiter die Bezeichnung

$$\bar{r}_A = \sum_{i \neq j}\sum r_{A_{ij}}\,\sigma_{A_i}\,\sigma_{A_j} : \sum_{i \neq j}\sum \sigma_{A_i}\,\sigma_{A_j} \tag{89}$$

eingeführt, so bekommt man endlich

$$\bar{r}_A = \frac{S^2{}_A - \sum\nolimits^{i}\sigma^2{}_{A_i}}{\left(\sum\nolimits^{i}\sigma_{A_i}\right)^2 - \sum\nolimits^{i}\sigma^2{}_{A_i}}. \tag{90}$$

Der Koeffizient $\bar{r}_A$ bildet folglich ein mit den Dispersionen gewogenes Mittel von den $\frac{n(n-1)}{2}$ Korrelationskoeffizienten, die zwischen den n ausgeglichenen Reihen gebildet werden können. Sind daher diese Koeffizienten sämtlich von ungefähr derselben Größe, und unterscheiden sich die Dispersionen nicht zu sehr voneinander, so muß $\bar{r}_A$ ein sehr zuverlässiges Maß für die durchschnittliche Stärke der Korrelationen liefern. Der Koeffizient ist in Analogie mit dem von BORTKIEWICZ[1]) definierten „Syndromiekoeffizienten“ abgeleitet, unterscheidet sich aber in seiner Bildungsweise von jenem dadurch, daß hier die verschiedenen Beobachtungszahlen der Reihen nirgends berücksichtigt worden sind.

[1]) Homogenität und Stabilität usw.

Wie der Syndromiekoeffizient leidet auch $\bar{r}_A$ an dem Fehler, daß die untere Grenze desselben für $n > 2$ nie $= -1$ werden kann, ein Umstand, der jedoch im folgenden seine Anwendbarkeit keineswegs einschränkt.

Der Koeffizient wird leicht aus (90) berechnet, nachdem gemäß (83), (85) und (87) die Hilfsgröße $S^2{}_A$ bestimmt worden ist.

Nach derselben Methode erhält man auch ohne Schwierigkeit ein Durchschnittsmaß für die Stärke der Korrelationen der zeitfreien Residualen. Nach (74) erhält man nämlich nach Summierung

$$\sum_{i \neq j}\sum r_{\xi_{ij}} \sigma_{\xi_i} \sigma_{\xi_j} = \frac{N}{N-u} \sum_{i \neq j}\sum r_{A_{ij}} \sigma_{A_i} \sigma_{A_j},$$

woraus man mit Hilfe von (88)

$$\sum_{i \neq j}\sum r_{\xi_{ij}} \sigma_{\xi_i} \sigma_{\xi_j} = \frac{N}{N-u} \left[S^2{}_A - \sum_i \sigma^2{}_{A_i} \right]$$

bekommt.

Wird nun analog zu (89)

$$\bar{r}_\xi = \sum_{i \neq j}\sum r_{\xi_{ij}} \sigma_{\xi_i} \sigma_{\xi_j} : \sum_{i \neq j}\sum \sigma_{\xi_i} \sigma_{\xi_j} \tag{91}$$

gesetzt, so erhält man

$$\bar{r}_\xi = \frac{N}{N-u} \times \frac{S^2{}_A - \sum_i \sigma^2{}_{A_i}}{\left(\sum_i \sigma_{\xi_i}\right)^2 - \sum_i \sigma^2{}_{\xi_i}}, \tag{92}$$

woraus der Koeffizient $\bar{r}_\xi$ mit Leichtigkeit bestimmt werden kann.

Derartig gewogene Koeffizienten werden auch gebraucht werden, um das Korrelationsverhältnis zwischen zwei Gruppen zu charakterisieren, von denen die eine aus n_1, die andere aus n_2 Reihen besteht. Sind $D_k^{(1)}$ und $D_k^{(2)}$ die für die betreffenden Gruppen nach (83) berechneten Werte der Größe D_k und

$$M = \frac{1}{N} \sum_k D_k^{(1)} D_k^{(2)}, \tag{93}$$

so ist, wie leicht ersichtlich,

$$M = \sum_i \sum_j r_{ij}^{(1,2)} \sigma_i^{(1)} \sigma_j^{(2)}, \tag{94}$$

wo

$r_{ij}^{(1,2)} =$ dem Korrelationskoeffizienten zwischen der i^{ten} Reihe der ersten und der j^{ten} Reihe der zweiten Gruppe,

$\sigma_i^{(1)} =$ der Dispersion der i^{ten} Reihe erster Gruppe,

$\sigma_i^{(2)} =$ „ „ „ „ „ zweiter „ .

Nach (67) hat man aber unter Anwendung leicht verständlicher Bezeichnungen

$$\sum i \sum j \, r_{ij}^{(1,2)} \sigma_i^{(1)} \sigma_j^{(2)} = \tag{95}$$

$$= \sum i \sum j \, r_{A_{ij}}^{(1,2)} \sigma_{A_i}^{(1)} \sigma_{A_j}^{(2)} + c_1 \sum i \, \beta_{i,1}^{\prime(1)} \sum i \, \beta_{i,1}^{\prime(2)} + c_2 \sum i \, \beta_{i,2}^{\prime(1)} \sum i \, \beta_{i,2}^{\prime(2)}.$$

Setzt man

$$M_A = M - c_1 \sum i \, \beta_{i,1}^{\prime(1)} \sum i \, \beta_{i,1}^{\prime(2)} - c_2 \sum i \, \beta_{i,2}^{\prime(1)} \sum i \, \beta_{i,2}^{\prime(2)}, \tag{96}$$

so bekommt man nach Einsetzung von (95) in (94)

$$\sum i \sum j \, r_{A_{ij}}^{(1,2)} \sigma_{A_i}^{(1)} \sigma_{A_j}^{(2)} = M_A, \tag{97}$$

und daraus

$$\gamma_A = \frac{M_A}{\sum i \, \sigma_{A_i}^{(1)} \sum i \, \sigma_{A_i}^{(2)}}, \tag{98}$$

wo der Koeffizient γ_A durch

$$\gamma_A = \sum i \sum j \, r_{A_{ij}}^{(1,2)} \sigma_{A_i}^{(1)} \sigma_{A_j}^{(2)} : \sum i \sum j \, \sigma_{A_i}^{(1)} \sigma_{A_j}^{(2)} \tag{99}$$

definiert ist.

Nach (74) ist weiter

$$\sum i \sum j \, r_{\xi_{ij}}^{(1,2)} \sigma_{\xi_i}^{(1)} \sigma_{\xi_j}^{(2)} = \frac{N}{N-u} \sum i \sum j \, r_{A_{ij}}^{(1,2)} \sigma_{A_i}^{(1)} \sigma_{A_j}^{(2)},$$

weshalb man also, falls

$$\gamma_\xi = \sum i \sum j \, r_{\xi_{ij}}^{(1,2)} \sigma_{\xi_i}^{(1)} \sigma_{\xi_j}^{(1)} : \sum i \sum j \, \sigma_{\xi_i}^{(1)} \sigma_{\xi_j}^{(2)}, \tag{100}$$

folgendes erhält:

$$\gamma_\xi = \frac{N}{N-u} \times \frac{M_A}{\sum i \, \sigma_{\xi_i}^{(1)} \sum i \, \sigma_{\xi_i}^{(2)}}. \tag{101}$$

Die Koeffizienten γ_A und γ_ξ sind also nach den resp. Definitionen, (99) und (100), gewogene Mittel der $n_1 \times n_2$ Korrelationskoeffizienten, die man bekommen würde, wenn man jede Reihe der einen Gruppe mit jeder Reihe der anderen korrelierte. γ_A und γ_ξ müssen daher, unter gleichen Voraussetzungen wie $\bar{r}_A$ und $\bar{r}_\xi$, durchschnittliche Maße für die Stärke der Korrelationen zwischen den Reihen der beiden Gruppen geben.

Nach Bestimmung der Hilfsgröße M_A nach (93) und (96) können γ_A und γ_ξ aus (98) und (101) leicht berechnet werden.

Für die Koeffizienten $\bar{r}_A$ und γ_A können selbstverständlich, wie für den Korrelationskoeffizienten r_A, gewisse Maximalwerte angegeben werden, die theoretisch nicht überschritten werden können. Wird in (74) $r_\xi = 1$ gesetzt, so bekommt man nach Summierung

$$\sum\sum_{i \neq j} \sigma_{\xi_i} \sigma_{\xi_j} = \frac{N}{N-u} \sum\sum_{i \neq j} r_A \sigma_{A_i} \sigma_{A_j}$$

und demnach gemäß (89)

$$\text{(102)} \qquad \text{Max.}\, \bar{r}_A = \frac{N-u}{N} \times \frac{\left(\sum_i \sigma_{\xi_i}\right)^2 - \sum_i \sigma^2_{\xi_i}}{\left(\sum_i \sigma_{A_i}\right)^2 - \sum_i \sigma^2_{A_i}}.$$

Dieses gibt die obere theoretische Grenze für $\bar{r}_A$ an. Eine untere Grenze kann nicht auf dieselbe Weise bestimmt werden, da bei $n > 2$ sämtliche Koeffizienten $r_{\xi_{ij}}$ niemals alle $= -1$ sein können. Dies zum Unterschiede vom Koeffizienten γ_A, für den offenbar sowohl eine obere als auch eine untere Grenze angegeben werden kann. Aus (74) bekommt man

$$\pm \sum_i \sum_j \sigma^{(1)}_{\xi_i} \sigma^{(2)}_{\xi_j} = \frac{N}{N-u} \sum_i \sum_j r^{(1,2)}_{A_{ij}} \sigma^{(1)}_{A_i} \sigma^{(2)}_{A_j},$$

und hieraus im Hinblick auf (99)

$$\text{(103)} \qquad \text{Max.}\, |\gamma_A| = \frac{N-u}{N} \times \frac{\sum_i \sigma^{(1)}_{\xi_i} \sum_i \sigma^{(2)}_{\xi_i}}{\sum_i \sigma^{(1)}_{A_i} \sum_i \sigma^{(2)}_{A_i}}.$$

Mit Hilfe dieser Formeln können nun für positive Werte der Koeffizienten die Zusammenhänge zwischen $\bar{r}_A$ und $\bar{r}_\xi$, γ_A und γ_ξ folgendermaßen geschrieben werden:

$$\bar{r}_A = \bar{r}_\xi \,\text{Max.}\bar{r}_A,$$
$$\gamma_A = \gamma_\xi \,\text{Max.}\gamma_A,$$

woraus hervorgeht, daß die Stärke der Korrelationen zwischen den zeitfreien Residualen ebenso gut aus einem Vergleich zwischen den direkt berechneten Koeffizienten und den Maximalwerten, wie aus den Koeffizienten $\bar{r}_\xi$ und γ_ξ selbst beurteilt werden kann. Im folgenden habe ich es meistens auch vorgezogen, neben $\bar{r}_A$ und γ_A die Maximalwerte statt der Korrelationskoeffizienten der zeitfreien Residualen anzugeben. Leztere würden für mehrere Altersgruppen größer als 1 ausfallen, und ich habe es daher nicht für zweckmäßig erachtet, sie anzuwenden.

Wie die Koeffizienten $\bar{r}_\xi$ und γ_ξ, sind auch die Formeln (102) und (103) von rein theoretischer Natur und gelten exakt nur für die erwartungsmäßigen Werte der einbegriffenen Größen. Würde daher bei Einsetzung der berechneten Werte $\bar{r}_A$ oder γ_A seinen Maximalwert übersteigen, so braucht dieses nicht notwendig einen Widerspruch zu enthalten, sondern der mittlere Fehler hat darüber zu entscheiden, ob die Abweichung von zufälliger Art ist oder nicht.

Aus den Formeln (102) und (103) können folgende Schlüsse gezogen werden. Wenn die Dispersionen normal sind in $n-1$ von den n Reihen, für die $\bar{r}_A$, und in allen Reihen in einer der beiden Gruppen, für die γ_A bestimmt worden ist, so müssen $\bar{r}_A$ und γ_A theoretisch $= 0$ sein, während mit abnehmender Stabilität die Möglichkeiten signifikativer Werte der Koeffizienten sich vergrößern. Umgekehrt gilt aber auch, daß signifikative Werte von $\bar{r}_A$ und γ_A übernormale Dispersionen in mindestens zwei Reihen notwendigerweise voraussetzen; betreffs γ_A in einer Reihe von jeder Gruppe. Da die Koeffizienten, wie es im folgenden der Fall ist, für 5jährige Gruppen berechnet werden, kann man außerdem berechtigt sein, auf durchschnittlich übernormale Dispersionen in allen Reihen zu schließen. Die Q^2_A-Werte einer Gruppe unterscheiden sich nämlich im allgemeinen nur unbedeutend voneinander. Als Hilfsmittel zur Beurteilung der Stabilität sind deshalb die Koeffizienten äußerst wertvoll. Um sie aber wirkungsvoll ausnutzen zu können, muß man die mittleren Fehler kennen. Diese habe ich unter der Voraussetzung bestimmt, daß die erwartungsmäßigen Werte sämtlicher Korrelationskoeffizienten, von denen $\bar{r}_A$ und γ_A gebildet sind, und somit auch die von $\bar{r}_A$ und γ_A selbst, $= 0$ sind. Man braucht nämlich nicht notwendigerweise Formeln, die für verschiedene Werte der Koeffizienten die entsprechenden mittleren Fehler liefern, sondern es genügt zu entscheiden, ob die erhaltenen Werte signifikativ sind oder nicht, und dafür sind diese mittleren Fehler völlig ausreichend. Die Bestimmung ist, wie folgt, durchgeführt worden. Wird

$$r_{A_{ij}}\,\sigma_{A_i}\,\sigma_{A_j} = m_{A_{ij}}$$

gesetzt, so ist nach (89)

$$\bar{r} = \sum\sum_{i \neq j} m_{A_{ij}} : \sum\sum_{i \neq j} \sigma_{A_i}\,\sigma_{A_j}$$

und somit, vorausgesetzt, daß die erwartungsmäßigen Werte von $m_{A_{ij}} = 0$ sind, in der ersten Approximation

$$\delta\bar{r}_A = \sum\sum_{i \neq j} \delta m_{A_{ij}} : \sum\sum_{i \neq j} \sigma_{A_i}\,\sigma_{A_j}. \tag{104}$$

Nach (78) ist nun für $m_{A_{ij}} = 0$

$$M(\delta^2 m_{A_{ij}}) = \frac{\sigma^2_{A_i}\,\sigma^2_{A_j}}{N-u},$$

und weiter kann ohne Schwierigkeit nachgewiesen werden, daß

$$M(\delta m_{A_{ij}}\,\delta m_{A_{hk}}) = 0, \text{ wenn } i \neq h \text{ oder } j \neq k \ (i \neq j).$$

Unter Berücksichtigung, daß $m_{A_{ij}} = m_{A_{ji}}$ ist, bekommt man daher aus (104) nach Quadrierung und Bildung des Durchschnittswertes

$$\varepsilon_0^2(\bar{r}_A) = \frac{2}{N-u} \sum\sum_{i \neq j} \sigma^2_{A_i}\,\sigma^2_{A_j} : \left(\sum\sum_{i \neq j} \sigma_{A_i}\,\sigma_{A_j}\right)^2. \tag{105}$$

In einer 5jährigen Gruppe unterscheiden sich im allgemeinen die σ_A-Werte nur wenig voneinander, und setzt man daher $\sigma_{A_i} = \sigma_{A_j}$, so ergibt sich

$$\varepsilon_0(r_A) = \sqrt{\frac{2}{n(n-1)(N-u)}}. \tag{106}$$

Auf ähnliche Weise und unter denselben Voraussetzungen kann leicht bewiesen werden, daß

$$\varepsilon_0{}^2(\gamma_A) = \frac{1}{N-u}\sum_i\sum_j \sigma^{2\,(1)}_{A_i}\sigma^{2\,(2)}_{A_j} : \left(\sum_i\sum_j \sigma^{(1)}_{A_i}\,\sigma^{(2)}_{A_j}\right)^2 \tag{107}$$

und daß demnach für $\sigma^{(1)}_{A_i} = \sigma^{(1)}_{A_j}$ und $\sigma^{(2)}_{A_i} = \sigma^{(2)}_{A_j}$

$$\varepsilon_0(\gamma_A) = \sqrt{\frac{1}{n_1 n_2 (N-u)}}. \tag{108}$$

Man beachte indessen, daß (106) und (108) Minimalwerte von (105) bzw. (107) sind und deswegen nur dann Geltung haben, wenn alle Dispersionen einer Gruppe von ungefähr derselben Größe sind. Im vorliegenden Falle sind sie für alle Gruppen fast exakt. Ein berechneter Wert von $\bar{r}_A$ oder γ_A, der den nach (106) und (108) bestimmten mittleren Fehler um mehr als das Dreifache übersteigt, kann daher als signifikativ betrachtet werden.

Die Korrelationen zwischen den Sterbenswahrscheinlichkeiten der verschiedenen Altersstufen in einer 5jährigen Gruppe sind jetzt mit Hilfe des Koeffizienten $\bar{r}_A$ und seines Maximalwertes Max. $\bar{r}_A$ untersucht worden. Für mehrere Altersgruppen konnte aber Max. $\bar{r}_A$ wegen des Auftretens imaginärer σ_ξ-Werte in einer oder in mehreren der einbegriffenen Reihen nicht bestimmt werden. Statt dessen habe ich hier den Maximalwert bestimmt, den man bekommen würde, wenn die Divergenzkoeffizienten in allen Reihen gleich groß und $= \bar{Q}^2{}_A$ wären. In solchem Falle wäre nach (76) für alle Werte von i und j

$$\text{Max. } r_{A_{ij}} = 1 - {}^1/_{\bar{Q}^2{}_A}$$

und demnach auch

$$\text{Max. } \bar{r}_A = 1 - {}^1/_{\bar{Q}^2{}_A}. \tag{109}$$

Diese Formel liefert Maximalwerte, die für Gruppen mit übernormalen Dispersionen in allen Reihen mit den nach (102) berechneten sehr nahe zusammenfallen, durchweg aber etwas höher ausfallen als jene. Aber auch in Gruppen mit unternormaler Dispersion in einer oder der anderen Reihe nehme ich an, daß Formel (109) ein einigermaßen zuverlässiges Maß für Max. $\bar{r}_A$ gibt. In einer 5jährigen Gruppe sind nämlich die Schwankungen der Divergenzkoeffizienten meistens von ganz zufälliger Natur, und die Stabilität der Gruppe kann daher durchschnittlich durch $\bar{Q}^2{}_A$ charakterisiert werden, auch wenn die Dispersionen in einer der Reihen oder in einigen unternormal wären.

In Tab. IX bezeichnen die unteren Ziffern in den Rechtecken $\bar{r}_A$ und die oberen Max. $\bar{r}_A$. Von den letzteren sind die eingeklammerten nach (109) und die übrigen nach (102) berechnet. Selbstverständlich kann Formel (109) nicht angewandt werden, wenn $\bar{Q}^2{}_A < 1$ ist.

Tab. IX.

Die Korrelation zwischen den Altersstufen einer Gruppe.

Die Werte von $\overline{r_A}$ und Max. $\overline{r_A}$.

Alter	Männer			Frauen		
	86—14	86—00	01—14	86—14	86—00	01—14
1— 5	0.923 + 0.773	0.939 + 0.825	0.828 + 0.683	0.899 + 0.673	0.915 + 0.714	0.825 + 0.663
5—10	0.727 + 0.694	0.766 + 0.708	0.576 + 0.693	0.774 + 0.708	0.781 + 0.793	0.741 + 0.783
15—20	0.366 + 0.387	0.410 + 0.333	(0.415) + 0.464	0.359 + 0.383	0.419 + 0.372	(0.213) + 0.229
25—30	0.259 + 0.292	0.348 + 0.292	(0.065) + 0.204	(0.123) + 0.139	(0.219) + 0.164	+ 0.028
35—40	0.313 + 0.245	(0.315) + 0.219	0.178 + 0.081	(0.123) + 0.191	(0.153) + 0.248	(0.074) + 0.189
45—50	0.259 + 0.287	0.329 + 0.406	0.144 — 0.010	0.277 + 0.193	0.422 + 0.266	+ 0.060
55—60	0.467 + 0.370	(0.660) + 0.389	(0.435) + 0.111	(0.342) + 0.303	0.399 + 0.446	(0.123) + 0.003
65—70	(0.500) + 0.449	(0.590) + 0.571	0.197 + 0.192	0.625 + 0.695	0.725 + 0.777	(0.160) + 0.303
75—80	0.616 + 0.512	0.632 + 0.697	0.497 + 0.080	0.751 + 0.710	0.804 + 0.781	0.538 + 0.525
	91—14	91—00	01—14	91—14	91—00	01—14
85—90	(0.394) + 0.285	(0.454) + 0.291	(0.296) + 0.202	0.493 + 0.526	0.597 + 0.699	(0.429) + 0.166

Nach (106) berechnet, bekommen die mittleren Fehler folgende Werte:

für die Periode 86—14 ..... $\varepsilon_0(\overline{r_A}) = 0.06$,

„ „ „ 86—00 und 01—14 ..... = 0.09.

Die Ausschließung der Altersstufe 0—1 von der ersten Gruppe ist dadurch veranlaßt worden, daß die Dispersion σ_A hier bedeutend größer ist als in den nächsten Altersstufen. Nähme man die Stufe 0—1 mit, so würden deshalb allen Koeffizienten $r_{A_{1j}}$ bei der Bestimmung von $\overline{r_A}$ entschieden größere Gewichte zugemessen werden als den übrigen. Folglich würde auch das Resultat gewissermaßen irreführend werden. Inwiefern diese Ausschließung auf die Größe der Koeffizienten einwirkt, wird später erörtert werden.

Aus den Zahlen der Tab. IX geht hervor, daß die Koeffizienten $\overline{r_A}$ in allen Altersgruppen, nur mit einer Ausnahme, positiv ausfallen, am größten in den Altersstufen 0—10 und in den Gruppen 65—70 und 75—80, am kleinsten in den mittleren Altersstufen. Sowohl für die Teilperiode 86—00, als auch für die ganze Zeitspanne können die Zahlen in fast allen Gruppen als signifikativ betrachtet werden, während dies für die Periode 01 bis 14 nur in den zehn ersten Altersstufen sowie in einzelnen von den folgenden Gruppen der Fall ist. Als Maß für die Stärke der Korrelationen bekommen aber diese Zahlen erst dann Bedeutung, wenn sie mit den entsprechenden Maximalwerten verglichen werden. Dabei stellt sich heraus, daß die r_A-Werte sich diesen in der Regel in allen Gruppen

sehr nahe anschließen, wo Korrelation mit einiger Sicherheit nachgewiesen werden kann. In mehreren Gruppen kommt es sogar vor, daß die Koeffizienten $\bar{r}_A$ ihre berechneten Maximalwerte übersteigen, die Abweichungen sind jedoch nicht größer, als daß sie ganz und gar der Unsicherheit der Zahlen zugeschrieben werden können. Im großen und ganzen kann man daher behaupten, daß, wo signifikative Werte vorliegen, die Koeffizienten auch fast ihre Maximalwerte erreicht haben oder, mit anderen Worten, die Korrelationen zwischen den zeitfreien Residualen sind beinahe vollständig. Dies muß dann bedeuten, daß die Faktoren, die die Sterblichkeit beeinflussen und die Schwankungen der Sterbenswahrscheinlichkeiten um die säkularen Regressionslinien verursachen, zum allergrößten Teil den verschiedenen Altersstufen einer Gruppe gemeinsam sind.

Die Zahlen der Tab. IX haben indessen auch ihre große Bedeutung als Stabilitätsmaße, und sie bestätigen in vorzüglicher Weise die Ergebnisse der vorhergehenden Stabilitätsuntersuchung. In fast allen Gruppen, wo man mit Hilfe der $\bar{Q}^2_A$-Werte auf durchschnittlich übernormale Dispersionen schließen konnte, sind nämlich die Koeffizienten $\bar{r}_A$ so groß, daß allein dieser Umstand genügen würde, um dieselben Folgerungen aufkommen zu lassen. Es kann daher keinem Zweifel unterliegen, daß die Übernormalität in diesen Gruppen, auch wo sie schwächer ausgeprägt ist, ihren Grund hat in äußeren störenden Einflüssen. Aber auch in mehreren solchen Gruppen, wo die niedrigen Q^2_A-Werte auf beinahe normale Dispersionen hindeuten, wie in den mittleren Altersgruppen während der Periode 01—14, scheint es, als ob die schwache Übernormalität nicht von ganz zufälliger Natur sein könne. Die Koeffizienten $\bar{r}_A$ fallen zwar im Verhältnis zu den mittleren Fehlern so gering aus, daß von der Stabilität nichts Bestimmtes gesagt werden kann — und so muß ja das Verhältnis der Theorie nach auch sein —, aber der Umstand, daß sie alle, nur mit einer Ausnahme, positiv sind, deutet doch darauf hin, daß sie nicht ganz und gar vom Zufall bestimmt sein können. Man wird daher sagen können, daß eine Tendenz zu übernormalen Dispersionen und positiven Korrelationen auch in diesen Gruppen vorherrscht; das wird im folgenden weiter bestätigt werden.

Der deutliche Parallelismus zwischen den Zahlen in den Tab. IX und IV ist natürlich eine Folge davon, daß die Koeffizienten $\bar{r}_A$ sich den Maximalwerten so nahe anschließen, welche ja ihrerseits von der Stabilität der Reihen abhängig sind. Infolgedessen werden die $\bar{r}_A$-Werte auch fast überall niedriger für die spätere Teilperiode, als für die nächst vorhergehende. Das Verhältnis zwischen $\bar{r}_A$ und Max. $\bar{r}_A$ scheint jedoch nicht größeren Veränderungen ausgesetzt gewesen zu sein. Freilich ist es während der Periode 01—14 wegen der niedrigen Werte der Koeffizienten und der damit verbundenen Unsicherheit der Zahlen sehr schwankend, aber im großen und ganzen muß man annehmen, daß die störenden Kräfte ungefähr ebenso stark korreliert sind, wie während der Periode vor 1900.

Die aus den Zahlen der Tab. IX gezogenen Schlüsse betreffs der korrelativen Eigenschaften der störenden Kräfte berechtigen nun zum Gebrauch des Koeffizienten γ_A bei der Untersuchung der Korrelation zwischen verschiedenen Gruppen. Dabei ist zunächst die Korrelation zwischen den Sterbenswahrscheinlichkeiten der Geschlechter betrachtet worden, und zwar so, daß aus jeder männlichen und entsprechenden weiblichen Gruppe der Koeffizient γ_A nach (98) bestimmt worden ist. Die Maximalwerte sind für Gruppen mit reellen σ_{ξ}-Werten nach (103), für die übrigen aus folgender Formel berechnet worden:

$$\text{Max. } \gamma_A = \sqrt{(1 - {}^1/_{\bar{Q}^2_{A_1}})(1 - {}^1/_{\bar{Q}^2_{A_2}})}. \tag{110}$$

Diese Formel bekommt man aus (98), wenn die Stabilität der einzelnen Gruppen durch $\overline{Q^2}_{A_1}$ und $\overline{Q^2}_{A_2}$ charakterisiert angenommen wird. Aus demselben Grunde wie vorher ist ferner die Altersstufe 0—1 für sich genommen worden, und die Stufen 1—5 bilden eine besondere Gruppe. Die Ergebnisse werden in Tab. X wiedergegeben, wo die unteren Ziffern die Werte von γ_A und die oberen die von Max. γ_A angeben. Die eingeklammerten Zahlen sind aus (110) berechnet.

Nach (106) bekommt man folgende Werte der mittleren Fehler:

für die Periode 86—14 ... $\varepsilon_0(\gamma_A) = 0.04$,
„ „ „ 86—00 u. 01—14 = 0.06.

Diese Zahlen zeigen hauptsächlich denselben Verlauf wie die $\overline{r}_A$-Werte der Tab. IX, und gleich wie jene schließen sie sich den berechneten Maximalwerten sehr wohl an, was hier auch für die meisten Gruppen der Periode 01—14 der Fall zu sein scheint. Die störenden Kräfte, die die Sterbenswahrscheinlichkeiten der Männer und Frauen in entsprechenden Altersgruppen beeinflussen, müssen daher, wo störende Kräfte überhaupt existieren, sehr stark positiv korreliert sein. Obgleich die Koeffizienten γ_A selbst, der erhöhten Stabilität zufolge, von der ersten bis zur zweiten Teilperiode abgenommen haben, scheinen dennoch die Korrelationen zwischen den störenden Kräften auch hier nicht abgeschwächt zu sein.

Als Stabilitätsmaße berechtigen diese Zahlen zu denselben Folgerungen, wie die $\overline{r}_A$-Werte der Tab. IX, und die durchgehenden positiven Werte der Koeffizienten bestätigen wiederum die Annahme, daß die schwache Übernormalität in den mittleren Altersstufen während der Periode 01—14 kaum von ganz zufälliger Art sein kann.

Tab. X.

Die Korrelation zwischen entsprechenden männlichen und weiblichen Altersgruppen.

Die Werte von γ_A und Max. γ_A.

Alter	86—14	86—00	01—14
0— 1	0.958 + 0.938	0.954 + 0.903	0.953 + 0.965
1— 5	0.913 + 0.766	0.928 + 0.809	0.832 + 0.660
5—10	0.755 + 0.756	0.780 + 0.776	0.660 + 0.754
15—20	0.369 + 0.253	0.484 + 0.197	(0.297) + 0.303
25—30	(0.189) + 0.224	(0.282) + 0.255	 + 0.138
35—40	(0.205) + 0.152	(0.220) + 0.164	(0.037) + 0.067
45—50	0.283 + 0.198	0.384 + 0.319	 + 0.026
55—60	(0.420) + 0.321	(0.554) + 0.419	(0.231) + 0.119
65—70	(0.560) + 0.558	(0.654) + 0.695	(0.212) + 0.161
75—80	0.682 + 0.603	0.715 + 0.698	0.528 + 0.381
	91—14	91—00	01—14
85—90	(0.491) + 0.480	(0.565) + 0.477	(0.356) + 0.298

Um die Zuverlässigkeit des Koeffizienten γ_A näher zu untersuchen, habe ich in den Gruppen 0—10, 35—40 und 75—80 die Korrelationskoeffizienten r_A zwischen den Sterbenswahrscheinlichkeiten entsprechender Altersstufen nach (67) berechnet. Die Ergebnisse gebe ich in Tab. XI wieder.

In den Altersstufen 0—10 und 75—80 halten sich die Zahlen, vom langsamen Abnehmen in der Gruppe 0—10 abgesehen, ziemlich konstant und werden durchschnittlich etwas größer als die für die resp. Gruppen berechneten γ_A-Werte. Das kann offenbar a priori vorausgesagt werden, aber besonders bemerkenswert ist es, daß der Unterschied nicht größer ist. Der Grund dafür liegt selbstverständlich in den starken Korrelationen zwischen den störenden Kräften innerhalb der einzelnen Gruppen. In der

Tab. XI.

Die Korrelation zwischen entsprechenden männlichen und weiblichen Altersstufen.

Die Werte des Koeffizienten r_A.

Alter	86—14	86—00	01—14
0— 1	+ 0.938	+ 0.908	+ 0.965
1— 2	+ 0.942	+ 0.933	+ 0.971
2— 3	+ 0.860	+ 0.880	+ 0.786
3— 4	+ 0.901	+ 0.906	+ 0.883
4— 5	+ 0.808	+ 0.905	+ 0.489
5— 6	+ 0.830	+ 0.830	+ 0.821
6— 7	+ 0.876	+ 0.909	+ 0.837
7— 8	+ 0.762	+ 0.787	+ 0.682
8— 9	+ 0.756	+ 0.754	+ 0.775
9—10	+ 0.571	+ 0.534	+ 0.643
35—36	+ 0.087	+ 0.363	— 0.290
36—37	+ 0.368	+ 0.643	— 0.053
37—38	+ 0.079	+ 0.075	+ 0.300
38—39	+ 0.289	+ 0.572	— 0.220
39—40	+ 0.097	+ 0.063	+ 0.249
75—76	+ 0.768	+ 0.828	+ 0.670
76—77	+ 0.642	+ 0.723	+ 0.330
77—78	+ 0.585	+ 0.744	+ 0.204
78—79	+ 0.621	+ 0.617	+ 0.564
79—80	+ 0.707	+ 0.658	+ 0.832

Gruppe 35—40 sind die Fluktuationen der Zahlen sehr viel größer, besonders während der beiden Teilperioden. Die störenden Kräfte sind aber hier so schwach, daß die ganz zufälligen Schwankungen der Sterbenswahrscheinlichkeiten dominieren, und man kann daher kaum eine größere Konstanz der Zahlen erwarten. Im Durchschnitt fallen sie jedoch ebenso groß aus, wie die entsprechenden r_A-Werte.

Es dürfte in diesem Zusammenhange von Interesse sein, darauf hinzuweisen, daß, obschon die r_A-Werte in den mittleren Altersstufen recht niedrig ausfallen, die direkt berechnete Korrelation zwischen den Sterblichkeitsziffern der ganzen männlichen und weiblichen Bevölkerung dennoch sehr stark ist, ja sogar alle bisher untersuchten Korrelationen übertrifft. Nach (67) berechnet, bekommen nämlich die Koeffizienten folgende Werte:

für die Periode 86—14 ... $r_A = + 0.972$,
„ „ „ 86—00 ..... $= + 0.980$,
„ „ „ 01—14 ..... $= + 0.879$.

Das erscheint vielleicht paradox, findet aber seine Erklärung natürlich zuerst in den äußerst starken Korrelationen zwischen den störenden Kräften, weiter aber auch darin, daß die Divergenzkoeffizienten und damit auch die Maximalwerte der Korrelationskoeffizienten bedeutend höher ausfallen, als in irgendeiner einzelnen Altersstufe.

Die Korrelationen zwischen den Sterbenswahrscheinlichkeiten der Geschlechter können in gewissen Punkten die Resultate erklären, die LEXIS[1]) und später u. a. GEIGEL[2]) und KAMMAN[3]) bei der Untersuchung der Stabilität des Geschlechtsverhältnisses der Gestorbenen gewonnen haben, in anderen Punkten scheinen sie jenen aber direkt zu widersprechen. So findet LEXIS in den Kinderjahren und den höchsten Altersstufen beinahe maximale Stabilität in der Quote, die das Verhältnis zwischen der Zahl gestorbener Männer und gestorbener Frauen angibt. In den mittleren Altersstufen treten dagegen stark übernormale Dispersionen hervor, und hier muß man deshalb, sagt LEXIS, „energisch wirkende und stark wechselnde äußere Ursachen annehmen, welche spezifisch auf die Sterblichkeit des einen oder des anderen Geschlechtes einwirken". Gesetzt, daß das Verhältnis zwischen den dem Sterben ausgesetzten männlichen und weiblichen Lebenden einigermaßen konstant ist, so ist natürlich die maximale Stabilität eine Folge der stark korrelativen Eigenschaften der Sterbenswahrscheinlichkeiten. Die übernormalen Dispersionen in den mittleren Altersstufen scheinen dagegen den hier erhaltenen Resultaten

[1]) Abhandlungen zur Theorie der Bevölkerungs- und Moralstatistik. Jena 1903.
[2]) Über die Stabilität des Geschlechtsverhältnisses der Gestorbenen. Freiburg i. B. 1880.
[3]) Das Geschlechtsverhältnis der Überlebenden in den Kinderjahren usw. Göttingen 1900.

bestimmt zu widerstreiten. Die störenden Kräfte haben sich nämlich erstens als ziemlich schwach herausgestellt, und zweitens sind sie positiv korreliert; man müßte daher statt dessen fast maximale Stabilität erwarten. Dies hat mich dazu veranlaßt, mit dem hier angewandten Material eine ähnliche Stabilitätsuntersuchung vorzunehmen. Für die Zeit 1901—1910 sind dabei nebenstehende Divergenzkoeffizienten erhalten worden.

Tab. XII.

Alter	Q^2
0— 1	0.69
0— 5	1.32
5—10	1.41
15—20	1.33
25—30	0.58
35—40	0.96
45—50	1.29
55—60	0.88
65—70	1.47
75—80	1.41

$\varepsilon_0(Q^2) = 0.47$

Im Gegensatz zu den LEXISschen Werten deuten diese auf annähernd normale Dispersionen in allen Altersgruppen und berechtigen also zu dem Schlusse, daß größere äußere Störungsursachen, die die Sterblichkeit des einen oder anderen Geschlechtes spezifisch beeinflußt haben, hier nicht vorgelegen haben.

Eine Untersuchung wie diese setzt indessen voraus, daß die relative Zahl lebender Männer und Frauen einigermaßen konstant ist. Das ist nun von LEXIS nicht befriedigend bewiesen worden, und GEIGEL hat es vollkommen übersehen. Wahrscheinlich können deshalb deren abweichende Resultate daraus erklärt werden, daß diese Bedingung nicht erfüllt gewesen ist.

Mit Hilfe des Koeffizienten γ_A habe ich weiter die Korrelationen zwischen verschiedenen Gruppen vom selben Geschlecht untersucht. Dabei bieten die Korrelationen zwischen den Gruppen 0—1, 1—5 und 5—10 besonderes Interesse und werden daher zuerst behandelt werden. Für jede der Kombinationen $(0—1) \times (1—5)$, $(0—1) \times (5—10)$ und $(1—5) \times (5—10)$ ist teils der Koeffizient γ_A nach (98), teils γ_ξ nach (101) berechnet worden. Der letztere wird hier infolge der stark übernormalen Dispersionen ziemlich sicher bestimmt und ist deshalb statt Max. γ_A gebraucht worden. Die Resultate sind in Tab. XIII zusammengestellt, wo die unteren Ziffern die γ_A-Werte, die oberen die γ_ξ-Werte angeben.

Aus diesen Zahlen geht hervor, daß auch die Korrelationen zwischen den störenden Kräften der verschiedenen Gruppen verhältnismäßig stark sind. Die Korrelation $(0—1) \times (1—5)$ ist bei Frauen schwächer als bei Männern, im übrigen scheinen aber entsprechende Korrelationen bei beiden Geschlechtern von ungefähr derselben Stärke zu sein. Durchgehend am höchsten sind die Koeffizienten für die naheliegenden Gruppen 1—5 und 5—10, etwas niedriger für 0—1 und 1—5 und am niedrigsten für 0—1 und 5—10. Daß die beiden letzten Korrelationen, besonders $(0—1) \times (5—10)$, die schwächsten sind, ist wohl die Folge davon, daß die Schwankungen der Säuglingssterblichkeit zum großen Teil auf ganz anderen Ursachen beruhen, als die Schwankungen der Sterbenswahrscheinlichkeiten der folgenden Altersstufen. Für die Kombinationen $(0—1) \times (1—5)$ und $(0—1) \times (5—10)$ nehmen die Koeffizienten auch in einer anderen Beziehung einen recht eigentümlichen Verlauf. Beide sind nämlich von der Periode vor bis zur Periode nach 1900 verschärft worden, und der letztere sehr beträchtlich. Das gilt nicht nur für die Koeffizienten γ_ξ, sondern auch für γ_A, die vorher fast ausnahmslos abnehmende Tendenz von der ersten bis zur zweiten Teilperiode gezeigt haben. Eine Erklärung dafür kann aber darin gefunden werden, daß entweder in den Altersstufen 0—1 oder 1—10 äußere Faktoren vorgekommen sind, die vor 1900 die Sterblichkeit in einer der Gruppen spezifisch beeinflußt haben, nach 1900 aber ganz oder teilweise verschwunden sind. Da nun die Stabilität

bei den 0—1jährigen von der ersten bis zur zweiten Periode keine größeren Veränderungen erlitten, in der Gruppe 1—5 sowie in 5—10 dagegen zugenommen hat, erscheint es glaubhaft, daß obengenannte Faktoren gerade in den letzten Altersgruppen wirksam gewesen sind. Eine große Rolle spielen sicher hierbei die Kinderkrankheiten, die gerade in diesen Altersstufen die Sterblichkeit am stärksten beeinflussen, und deren Wirkungen, wie vorhin erwähnt, nach 1900 bedeutend abgeschwächt worden sind.

Tab. XIII.

Die Korrelation zwischen verschiedenen Gruppen im Alter von 0 bis 10 Jahren.

Werte der Koeffizienten γ_A und γ_ξ.

Männer.

1886—1914.

Alter	0—1	1—5
1— 5	+ 0.642 + 0.606	
5—10	+ 0.478 + 0.401	+ 0.752 + 0.619

1886—1900.

Alter	0—1	1—5
1— 5	+ 0.616 + 0.586	
5—10	+ 0.372 + 0.320	+ 0.772 + 0.657

1901—1914.

Alter	0—1	1—5
1— 5	+ 0.701 + 0.629	
5—10	+ 0.657 + 0.491	+ 0.718 + 0.501

Frauen.

1886—1914.

Alter	0—1	1—5
1— 5	+ 0.519 + 0.481	
5—10	+ 0.446 + 0.384	+ 0.754 + 0.632

1886—1900.

Alter	0—1	1—5
1— 5	+ 0.501 + 0.467	
5—10	+ 0.351 + 0.304	+ 0.771 + 0.656

1901—1914.

Alter	0—1	1—5
1— 5	+ 0.538 + 0.478	
5—10	+ 0.607 + 0.513	+ 0.765 + 0.600

Ein Vergleich zwischen den $\bar{r}_A$-Werten der Gruppe 1—5, Tab. IX, und den γ_A-Werten für die Kombination (0—1) × (1—5) zeigt ferner, daß die Korrelationen innerhalb der Gruppe 1—5 durchschnittlich höher sind, als die Korrelationen zwischen den Sterbenswahrscheinlichkeiten der 0—1- und 1—5jährigen, und zwar, wie oben bemerkt wurde, vermutlich deswegen, weil die Sterblichkeit in diesen beiden Gruppen von teilweise ungleichen Faktoren beeinflußt wird. Die Mitnahme der Altersstufe 0—1 bei der Berechnung von $\bar{r}_A$ würde deshalb einen kleineren Wert dieses Koeffizienten verursachen.

Von der Gruppe 5—10 an sind ferner die Korrelationen zwischen naheliegenden Gruppen mittels γ_A und Max. γ_A untersucht worden. Die Resultate sind in Tab. XIV wiedergegeben, wo die oberen Zahlen wieder die Maximalwerte angeben. Die eingeklammerten Zahlen sind nach (110) berechnet. Die Korrelationen zwischen den Gruppen 75—80 und 85—90 sind nur für die Periode 01—14 bestimmt worden, da die Reihen vor 1900 in den beiden Gruppen nicht gleich lang sind.

Tab. XIV.

Die Korrelation zwischen naheliegenden Gruppen.

Die Werte von γ_A und Max. γ_A.

Korrelation	Männer			Frauen		
	86—14	86—00	01—14	86—14	86—00	01—14
(5—10) × × (15—20)	0.599 + 0.434	0.566 + 0.397	(0.515) + 0.486	0.465 + 0.220	0.578 0.153	(0.399) + 0.329
(15—20) × × (25—30)	0.315 + 0.231	0.371 + 0.228	(0.164) + 0.190	(0.218) + 0.198	(0.312) + 0.156	+ 0.140
(25—30) × × (35—40)	0.291 + 0.224	(0.338) + 0.235	(0.112) + 0.140	(0.123) + 0.078	(0.183) + 0.121	+ 0.027
(35—40) × × (45—50)	0.300 + 0.292	(0.360) + 0.318	0.173 + 0.151	(0.197) + 0.202	(0.263) + 0.247	+ 0.190
(45—50) × × (55—60)	0.366 + 0.249	(0.521) + 0.349	(0.317) + 0.041	(0.328) + 0.303	0.419 + 0.401	+ 0.095
(55—60) × × (65—70)	(0.507) + 0.462	(0.624) + 0.544	(0.350) + 0.215	(0.463) + 0.436	0.545 + 0.583	(0.140) + 0.144
(65—70) × × (75—80)	(0.559) + 0.516	(0.622) + 0.647	0.329 + 0.215	0.685 + 0.643	0.763 + 0.738	(0.303) + 0.342
(75—80) × × (85—90)			(0.395) + 0.250			(0.496) + 0.390

Die mittleren Fehler sind hier dieselben, wie für Tab. X, d. h.

für die Periode 86—14 ..... $\varepsilon_0(\gamma_A) = 0.04$,
„ „ „ 86—00 und 01—14 ..... $= 0.06$.

Der Verlauf dieser Zahlen ist in der Hauptsache derselbe wie bei den $\overline{r}_A$- und γ_A-Werten in den Tab. IX und X. Wie diese sind die Koeffizienten im Verhältnis zu den Maximalwerten sehr hoch, zeigen also an, daß auch hier die Korrelationen zwischen den zeitfreien Residualen beinahe vollständig sind, oder mit anderen Worten, daß die störenden Kräfte naheliegender Gruppen sehr stark positiv korreliert sind. Die aus den vorhergehenden Tabellen betreffs der Stabilität gewonnenen Schlußsätze findet man hier wieder bestätigt, und die durchgehenden positiven Werte der Koeffizienten scheinen eine schwache Übernormalität auch in den mittleren Altersstufen während der Periode 01—14 anzudeuten.

Es würde zu mühevoll werden, für eine fortgesetzte Betrachtung der Korrelationen zwischen verschiedenen Gruppen den Koeffizienten γ_A anzuwenden, und zwar der vielen Kombinationen wegen, für welche in solchem Falle die γ_A-Werte berechnet werden müssen. Ich habe es deshalb nochmals vorgezogen, mich einer mehr summarischen Behandlungsweise zu bedienen. Hierbei habe ich die Altersgruppen zu drei größeren Gruppenordnungen — I, II und III — vereinigt, so daß

zu I die Gruppen 0—1, 1—5 und 5—10 gehören,
„ II „ „ 15—20, 25—30, 35—40 und 45—50 gehören,
„ III „ „ 55—60, 65—70 und 75—80 gehören.

Für diese Gruppenordnungen habe ich dann $\bar{r}_A$ und γ_A entsprechende gewogene Koeffizienten berechnet. Die Gruppe 85—90 konnte nicht in III mitgenommen werden, da die Reihen nicht gleich lang sind, wie in den übrigen Gruppen.

Die Dispersion in der i^{ten} ausgeglichenen Reihe der p^{ten} Gruppe einer Gruppenordnung sei mit $\sigma_{A_i}^{(p)}$ bezeichnet, der Korrelationskoeffizient zwischen der i^{ten} ausgeglichenen Reihe der p^{ten} Gruppe und der j^{ten} ausgeglichenen Reihe der q^{ten} Gruppe mit $r_{A_{ij}}^{(pq)}$. Weiter sei

$$M_{A_{pq}} = \sum_i \sum_j r_{A_{ij}}^{(pq)} \sigma_{A_i}^{(p)} \sigma_{A_j}^{(q)} \tag{111}$$

und

$$K_A = \sum_{p \neq q} \sum \varkappa_p \varkappa_q M_{A_{pq}}, \tag{112}$$

wo $\varkappa_p$ und $\varkappa_q$ zweckmäßig gewählte Gewichte sind. Bildet man nun

$$\bar{r}_A' = \frac{K_A}{\sum_{p \neq q} \sum \varkappa_p \varkappa_q \sum_i \sum_j \sigma_{A_i}^{(p)} \sigma_{A_j}^{(q)}}, \tag{113}$$

so ist $\bar{r}_A'$ ein gewogenes Mittel von sämtlichen Korrelationskoeffizienten, die innerhalb der Gruppenordnung zwischen Reihen verschiedener Gruppen gebildet werden können. Dagegen werden die vorhin untersuchten Korrelationen zwischen den Reihen, die zu derselben Gruppe gehören, nicht berücksichtigt. Dieser Koeffizient gibt daher ein Durchschnittsmaß für die Stärke der Korrelationen zwischen weiter voneinander entfernt liegenden Reihen innerhalb der Gruppenordnung.

Die Größe $\varkappa_p$ soll natürlich so gewählt werden, daß sämtlichen Korrelationskoeffizienten, aus welchen $\bar{r}_A'$ gebildet wird, ungefähr gleich große Gewichte zugeteilt werden. Vorausgesetzt, daß die Dispersionen der verschiedenen Reihen einer Gruppe sich nicht allzusehr voneinander unterscheiden, erreicht man dieses, wenn

$$\varkappa_p = 1 : \frac{1}{n_p} \sum_i \sigma_{A_i}^{(p)} \tag{114}$$

gesetzt wird, wobei $n_p =$ der Zahl der Reihen in der p^{ten} Gruppe ist.

Die Formel (113) kann dann auch so geschrieben werden:

$$\bar{r}_A' = \frac{K_A}{\left(\sum_p n_p\right)^2 - \sum_p n^2{}_p}. \tag{115}$$

Für die numerische Berechnung von $\bar{r}_A'$ kann der Ausdruck für K_A folgendermaßen reduziert werden. Nach (96) hat man

$$M_{A_{pq}} = M_{pq} - c_1 \sum_i \beta_{i,1}^{\prime(p)} \sum_i \beta_{i,1}^{\prime(q)} - c_2 \sum_i \beta_{i,2}^{\prime(p)} \sum_i \beta_{i,2}^{\prime(q)},$$

und somit

$$(116)\qquad K_A = \sum_{p \neq q}\sum x_p x_q M_{pq} - c_1\left(\sum_p x_p \sum_i \beta'^{(p)}_{i,1}\right)^2 + c_1 \sum_p x^2_p \left(\sum_i \beta'^{(p)}_{i,1}\right)^2 -$$

$$- c_2\left(\sum_p x_p \sum_i \beta'^{(p)}_{i,2}\right)^2 + c_2 \sum_p x^2_p \left(\sum_i \beta'^{(p)}_{i,2}\right)^2.$$

Nun ist nach (93)

$$M_{pq} = \frac{1}{N}\sum_k D^{(p)}_k D^{(q)}_k,$$

weshalb man also

$$\sum_{p \neq q}\sum x_p x_q M_{pq} = \frac{1}{N}\sum_k \left(\sum_p x_p D^{(p)}_k\right)^2 - \frac{1}{N}\sum_k \sum_p x^2_p D^{2(p)}_k$$

erhält.

Setzt man hier

$$(117)\qquad \Delta_k = \sum_p x_p D^{(p)}_k$$

und beachtet, daß nach (85)

$$\frac{1}{N}\sum_k D^{2(p)}_k = S^2_p$$

und ferner, daß nach (87)

$$S^2_{A_p} = S^2_p - c_1\left(\sum_i \beta'^{(p)}_{i,1}\right)^2 - c_2\left(\sum_i \beta'^{(p)}_{i,2}\right)^2$$

ist, so bekommt man nach Einsetzung in (116)

$$(118)\qquad K_A = \frac{1}{N}\sum_k \Delta^2_k - \sum_p x^2_p S^2_{A_p} - c_1\left(\sum_p x_p \sum_i \beta'^{(p)}_{i,1}\right)^2 - c_2\left(\sum_p x_p \sum_i \beta'^{(p)}_{i,2}\right)^2.$$

Diese Formel enthält außer der Größe Δ_k nur bereits bekannte Ausdrücke, und K_A kann deshalb daraus mit Leichtigkeit bestimmt werden, nachdem Δ_k nach (117) berechnet worden ist.

Den theoretischen Maximalwert des Koeffizienten $\bar{r}'_A$ habe ich nur unter der Voraussetzung bestimmt, daß die Stabilität einer Gruppe durchschnittlich durch $\bar{Q}^2_A$ charakterisiert werden kann, so daß für alle Werte von i und j

$$\text{Max. } r^{(pq)}_{A_{ij}} = \sqrt{\left(1 - 1/\bar{Q}^2_{A_p}\right)\left(1 - 1/\bar{Q}^2_{A_q}\right)}$$

ist. Man bekommt dann aus (115)

$$(119)\qquad \text{Max. } \bar{r}'_A = \frac{\left(\sum_p n_p \sqrt{1 - 1/\bar{Q}^2_{A_p}}\right)^2 - \sum_p n^2_p \left(1 - 1/\bar{Q}^2_{A_p}\right)}{\left(\sum_p n_p\right)^2 - \sum_p n^2_p}.$$

Vorausgesetzt, daß die σ_A-Werte innerhalb einer Gruppe sich nicht zu sehr voneinander unterscheiden, kann weiter nach derselben Methode, welche zur Bestimmung von $\varepsilon_0(r_A)$ angewandt wurde, gezeigt werden, daß

$$\varepsilon_0(\bar{r}'_A) = \sqrt{\frac{2}{(N-u)\left[\left(\sum^p n_p\right)^2 - \sum^p n^2{}_p\right]}}. \tag{120}$$

Auch die Korrelationen zwischen verschiedenen Gruppenordnungen habe ich mittels ähnlich gewogener Koeffizienten untersucht. $r^{(p_1 p_2)}_{A_{ij}}$ bezeichne den Korrelationskoeffizienten zwischen der i^{ten} ausgeglichenen Reihe der $p_1{}^{\text{ten}}$ Gruppe in der einen Gruppenordnung und der j^{ten} ausgeglichenen Reihe der $p_2{}^{\text{ten}}$ Gruppe in der anderen, $\sigma^{(p_1)}_{A_i}$ und $\sigma^{(p_2)}_{A_j}$ seien die betreffenden Dispersionen, und

$$M_{A_{p_1 p_2}} = \sum i \sum j \; r^{(p_1 p_2)}_{A_{ij}} \sigma^{(p_1)}_{A_i} \sigma^{(p_2)}_{A_j}.$$

Wird hiervon

$$L_A = \sum^{p_1} \sum^{p_2} \varkappa_{p_1} \varkappa_{p_2} M_{A_{p_1 p_2}}$$

gebildet und weiter

$$\gamma'_A = \frac{L_A}{\sum^{p_1} \sum^{p_2} \varkappa_{p_1} \varkappa_{p_2} \sum i \sum j \; \sigma^{(p_1)}_{A_i} \sigma^{(p_2)}_{A_j}},$$

so ist der Koeffizient γ'_A ein gewogenes Mittel von sämtlichen Korrelationskoeffizienten, die zwischen den Reihen der beiden Gruppenordnungen gebildet werden können. Setzt man wie vorher

$$\varkappa_p = 1 : \frac{1}{n_p} \sum i \; \sigma^{(p)}_{A_i},$$

so kann der Ausdruck für γ'_A auch so geschrieben werden:

$$\gamma'_A = \frac{L_A}{\sum^{p_1} n_{p_1} \sum^{p_2} n_{p_2}}. \tag{121}$$

Die Größe L_A erhält man hier folgendermaßen. Nach (96) ist

$$M_{A_{p_1 p_2}} = M_{p_1 p_2} - c_1 \sum i \; \beta'^{(p_1)}_{i,1} \sum i \; \beta'^{(p_2)}_{i,1} - c_2 \sum i \; \beta'^{(p_1)}_{i,2} \sum i \; \beta'^{(p_2)}_{i,2}$$

und demnach

$$L_A = \sum^{p_1} \sum^{p_2} \varkappa_{p_1} \varkappa_{p_2} M_{p_1 p_2} - c_1 \sum^{p_1} \varkappa_{p_1} \sum i \; \beta'^{(p_1)}_{i,1} \times \sum^{p_2} \varkappa_{p_2} \sum i \; \beta'^{(p_2)}_{i,1} - \tag{122}$$

$$- c_2 \sum^{p_1} \varkappa_{p_1} \sum i \; \beta'^{(p_1)}_{i,2} \times \sum^{p_2} \varkappa_{p_2} \sum i \; \beta'^{(p_2)}_{i,2}.$$

Weiter hat man nach (93)

$$M_{p_1 p_2} = \frac{1}{N} \sum_k D_k^{(p_1)} D_k^{(p_2)}$$

und sodann, wenn nach (117) die Bezeichnungen

$$\Delta_k^{(1)} = \sum_p x_{p_1} D_k^{(p_1)} \text{ und } \Delta_k^{(2)} = \sum_{p_2} x_{p_2} D_k^{(p_2)}$$

angewandt werden,

$$\sum_{p_1} \sum_{p_2} x_{p_1} x_{p_2} M_{p_1 p_2} = \frac{1}{N} \sum_k \Delta_k^{(1)} \Delta_k^{(2)}.$$

Mit Hilfe hiervon kann die Größe L_A ohne Schwierigkeit aus (122) berechnet werden.

Unter denselben Voraussetzungen, wie vorher, kann weiter nachgewiesen werden, daß

$$\text{Max. } \gamma_A' = \frac{\sum_{p_1} n_{p_1} \sqrt{1 - 1/\overline{Q}^2_{A p_1}} \sum_{p_2} n_{p_2} \sqrt{1 - 1/\overline{Q}^2_{A p_2}}}{\sum_{p_1} n_{p_1} \sum_{p_2} n_{p_2}} \tag{123}$$

und daß

$$\varepsilon_0(\gamma_A') = \sqrt{\frac{1}{(N-u) \sum_{p_1} n_{p_1} \sum_{p_2} n_{p_2}}}. \tag{124}$$

Obgleich die Untersuchungen mit Hilfe dieser Koeffizienten sehr vereinfacht werden, erheischen sie doch nicht unbedeutende numerische Rechnungen, und ich habe deshalb die $\overline{r}_A'$- und γ_A'-Werte nur für den ganzen Zeitraum 1886—1914 berechnet.

In Tab. XV geben die unteren Zahlen in jedem Felde die $\overline{r}_A'$-Werte und die oberen die Max. $\overline{r}_A'$-Werte, nach Formel (115) bzw. (119) berechnet. Die mittleren Fehler sind hier nach (120)

für Gruppenordnung I ..... $\varepsilon_0(\overline{r}_A') = 0.04$,
„ „ II und III ..... $= 0.02$.

Die Korrelationen in der ersten Gruppenordnung sind früher mittels γ_A und γ_ξ untersucht worden, und die hier berechneten Zahlen, die ja übrigens ganz summarischer Art sind, bringen daher keine neuen Resultate zutage. In den Gruppenordnungen II und III sind die $\overline{r}_A'$-Werte ungefähr ebenso groß, wie die vorher für naheliegende Gruppen innerhalb der einzelnen Gruppenordnungen berechneten γ_A-Koeffizienten (Tab. XIV). Das Verhältnis zwischen $\overline{r}_A'$ und Max. $\overline{r}_A'$ liegt aber hier etwas tiefer, ungefähr 0.7—0.9, deutet aber dennoch

Tab. XV.

Die Werte von $\overline{r}_A'$ und Max. $\overline{r}_A'$.

Gruppenordnung	Männer	Frauen
I	0.846 + 0.549	0.857 + 0.521
II	0.353 + 0.268	0.217 + 0.147
III	0.545 + 0.475	0.553 + 0.493

wieder auf starke korrelative Eigenschaften bei den störenden Kräften hin, und zwar auf etwas stärkere in der III. als in der II. Gruppenordnung. Auch die Schwankungen in den Sterbenswahrscheinlichkeiten weiter voneinander entfernt liegender Altersstufen müssen also zum großen Teil gemeinsamen äußeren Störungsursachen zugeschrieben werden.

Um nochmals die unsichereren Korrelationsverhältnisse in den mittleren Altersstufen in der Periode 01—14 zu beleuchten, habe ich die Gruppen zusammengeschoben, die die kleinsten $\bar{Q}^2_A$-Werte aufweisen, nämlich die Gruppen 25—30, 35—40 und 45—50, und für sie folgende $\bar{r}'_A$-Werte gefunden:

für Männer .... $\bar{r}'_A = + 0.103$
„ Frauen ..... $= + 0.078$ $\quad \varepsilon_0(\bar{r}'_A) = 0.082.$

Mit Bestimmtheit können also auch nicht aus diesen Ziffern irgendwelche Schlüsse gezogen werden, weder über Korrelation noch über Stabilität, aber wie vorher deuten sie auf schwache, positiv korrelierte, störende Kräfte.

Für die Zeit 86—14 sind weiter die Korrelationen zwischen verschiedenen Gruppenordnungen mit Hilfe der Koeffizienten γ'_A und Max. γ'_A, (121) und (123), untersucht worden. Infolge der soeben nachgewiesenen starken Korrelationen zwischen den störenden Kräften innerhalb der betreffenden Gruppenordnungen müssen diese Koeffizienten trotz ihrem summarischen Charakter recht zuverlässige Korrelationsmaße sein. Die Koeffizienten sind für die drei Kombinationen berechnet worden, die von den Gruppenordnungen gebildet werden können, und die Ergebnisse sind in Tab. XVI zusammengestellt, wo wie gewöhnlich die oberen Zahlen die Maximalwerte angeben.

Tab. XVI.

Die Werte von γ'_A und Max. γ'_A.

Männer.

Gruppenordnung	I	II
II	0.542 + 0.364	
III	0.674 + 0.305	0.439 + 0.257

Frauen.

Gruppenordnung	I	II
II	0.434 + 0.174	
III	0.689 + 0.331	0.353 + 0.230

Die mittleren Fehler, nach (124) berechnet, sind für keine Kombination größer als 0.02.

Das Verhältnis zwischen γ'_A und Max. γ'_A variiert hier zwischen 0.4 und 0.7, ohne jedoch einen bestimmten Verlauf zu nehmen. Mit den Maximalwerten verglichen, fallen also diese Koeffizienten niedriger aus, als alle bisher untersuchten, deuten aber noch auf ziemlich starke positive Korrelationen zwischen den störenden Kräften hin. Die Schwankungen der Sterbenswahrscheinlichkeiten um die säkularen Regressionslinien scheinen also in allen Altersstufen zum großen Teil denselben äußeren Störungsursachen zugeschrieben werden zu können. Das wird man weiter bestätigt finden, wenn sämtliche Gruppen vereinigt werden, und darauf der Koeffizient $\bar{r}'_A$ für die ganze Gruppenordnung nach (115) berechnet wird. Dabei erhält man nämlich

für Männer ..... $\bar{r}'_A = + 0.317$ $\quad$ Max. $\bar{r}'_A = 0.526,$
„ Frauen ..... $\bar{r}'_A = + 0.246$ $\quad$ Max. $\bar{r}'_A = 0.454.$

Diese Zahlen, die von nicht weniger als 904 verschiedenen Korrelationskoeffizienten die gewogenen Mittel bilden, sind nicht nur absolut, sondern auch relativ im Verhältnis zu den Maximalwerten von fast derselben Größe wie die γ'_A-Koeffizienten der Tab. XVI und bestätigen also die daraus gewonnenen Schlußsätze. Selbstverständlich schließt dieses nicht aus, daß andere Störungsursachen auf gewisse Altersgrenzen beschränkt sein können. So wird zweifelsohne die Säuglingssterblichkeit zum großen Teil von ihren speziellen Faktoren verursacht, und die großen Influenzaepidemien in den 90er Jahren hatten fast nur in den mittleren Altersstufen auf die Sterblichkeit Einfluß.

Endlich sind mittels der Koeffizienten γ'_A und Max. γ'_A auch die Korrelationen zwischen den Sterbenswahrscheinlichkeiten des männlichen und weiblichen Geschlechts für die verschiedenen Gruppenordnungen untersucht worden. Auf Grund der zuletzt gewonnenen Resultate und infolge der früher nachgewiesenen starken Korrelationen zwischen den störenden Kräften in entsprechenden Gruppen kann man erwarten, daß die Fluktuationen der Sterbenswahrscheinlichkeiten der Geschlechter auch für größere Gruppen zum großen Teil gemeinsamen Störungsursachen zugeschrieben werden können. Daß dies auch der Fall ist, wird durch die Zahlen der Tab. XVII bestätigt. Das Verhältnis zwischen γ'_A und Max. γ'_A stellt sich nämlich zwar als geringer heraus, als durchschnittlich für die verschiedenen Gruppen der betreffenden Gruppenordnungen, ist aber doch ziemlich hoch, ca. 0.7—0.9, am kleinsten in der II. und am größten in der III. Gruppenordnung. Für die in der Tabelle zuletzt angegebenen Zahlen, welche nach Vereinigung sämtlicher männlichen und weiblichen Gruppen berechnet worden sind, ist das Verhältnis γ'_A : Max. γ'_A ungefähr $= 0.6$, d. h. etwas kleiner als für die verschiedenen Gruppenordnungen, was ja auch zu erwarten ist. Die Übereinstimmung muß aber als ziemlich gut betrachtet werden, und die Zahlen lassen die Behauptung berechtigt erscheinen, daß die Schwankungen der Sterbenswahrscheinlichkeiten des männlichen und weiblichen Geschlechts in allen Altersstufen meistens von denselben störenden Faktoren verursacht werden müssen.

Tab. XVII.

Gruppenordnung	Max. γ'_A und γ'_A
I	0.840 + 0.636
II	0.280 + 0.189
III	0.553 + 0.477
I—III	0.473 + 0.267

$\varepsilon_0(\gamma'_A) < 0.02$

Obgleich natürlich noch mancherlei von Interesse durch ein fortgesetztes und eingehenderes Studium der Korrelationen hervorzuheben wäre, will ich dennoch diesen Abschnitt hier abschließen. Aus räumlichen und zeitlichen Gründen konnten mehrere wichtige Probleme, wie z. B. die Frage von „the natural selection in man“, nicht einmal gestreift werden, und andere wurden nur ganz flüchtig behandelt. Es war aber nicht die Absicht, eine erschöpfende Behandlung aller hierher gehörenden Fragen zu geben — das würde jahrelange Arbeit erfordern —, sondern nur eine orientierende Vorstellung von den korrelativen Eigenschaften der Sterbenswahrscheinlichkeiten in verschiedenen Altersstufen und Altersgruppen.

Kap. VI.

Die Effektivität der Ausgleichung. Die Einwirkung der „Seuchensterblichkeit" auf Stabilität und Korrelation.

Bisher ist stillschweigend angenommen worden, daß die in den vorhergehenden Kapiteln behandelten Resultate vollständig oder wenigstens annähernd vom Einflusse der säkularen Variation befreit waren, und sie sind natürlich nur unter dieser Voraussetzung vollgültig. Daß diese Bedingung sehr genau erfüllt werden muß, ergibt sich bei den Konstruktionen der säkularen Regressionslinien, aber um die Zuverlässigkeit der Resultate auch numerisch zu untersuchen, habe ich für die Zeitspanne 1886—1914 die Ausgleichung in einigen Altersstufen bis zum 4. Grad betrieben, d. h. auch auf die Konstanten β_3' und β_4' Rücksicht genommen. Offenbar wird bei fortgesetzter Ausgleichung die Dispersion σ_A kleiner für jedes neue Glied, das in der Gleichung der Regressionslinie mitgenommen wird, um schließlich $= 0$ zu werden, wenn nicht früher, so jedenfalls, wenn die Zahl der Unbekannten $= N$ wird. In dem Begriffe „säkulare Variation" liegt indessen, daß nur die großen säkularen Schwankungen eliminiert werden sollen, und daß deshalb die Ausgleichung mit möglichst wenigen Gliedern vorgenommen werden muß. Ergibt sich daher, daß die Mitnahme der β_3'- und β_4'-Konstanten die vorhergehenden Resultate nicht in höherem Grade verändert, so zeigt dies, daß die säkulare Variation mit einer Kurve 2. Grades vollkommen befriedigend eliminiert werden kann, wenigstens, wenn die Reihen nicht länger sind als die hier untersuchten. Wenn die β_3'- und β_4'-Konstanten sehr klein sind, können die Divergenzkoeffizienten, wie unmittelbar aus (54) ersichtlich ist, sogar zunehmende Tendenz zeigen. Dann hat man offenbar dem Minimalwert den Vorzug zu geben.

Ehe ich aber die Resultate wiedergebe, möchte ich zunächst als Beispiel für die Einfachheit der in Kap. II beschriebenen Methode für die männliche Altersstufe 0—1 eine vollständige Bestimmung der β'-Konstanten und der Dispersionen der ausgeglichenen Reihen durchführen.

Bei $N = 29$ geben die Formelsysteme (39) und (39*)

$$\begin{aligned} \alpha_{20} &= -\ 70, & c_1 &= \ \ 70, \\ \alpha_{31} &= -125.8, & c_2 &= 3906, \\ \alpha_{42} &= -179.29, & c_3 &= 2089 \times 10^2, \\ \alpha_{40} &= \ \ 3744, & c_4 &= 1095 \times 10^4. \end{aligned}$$

Die Bestimmung der d-Momente erfolgt am besten nach CHARLIERs[1]) selbstkontrollierendem Rechenschema.

---

[1]) Siehe Vorlesungen über die Grundzüge usw.

Tab. XVIII.

$N = 29.$ $m = 10.0.$

$k - \frac{N+1}{2} = x$	F	$F - m$	$(F-m)^2$	$x(F-m)$	$x^2(F-m)$	$x^3(F-m)$	$x^4(F-m)$	$(x+1)^4(F-m)$
— 14	12.197	+ 2.197	4.827	— 30.758	+ 430.612	— 6 028.568	+ 84 399.952	+ 62 748.517
— 13	11.209	+ 1.209	1.462	— 15.717	+ 204.321	— 2 656.173	+ 34 530.249	+ 25 069.824
— 1	12.055	+ 2.055	4.223	— 2.055	+ 2.055	— 2.055	+ 2.055	0
0	10.923	+ 0.923	0.852	0	0	0	0	+ 0.923
+ 1	11.235	+ 1.235	1.525	+ 1.235	+ 1.235	+ 1.235	+ 1.235	+ 19.760
+ 14	8.061	— 1.939	3.760	— 27.146	— 380.044	— 5 320.616	— 74 488.624	— 98 161.875
Σ		+ 1.940	60.206	— 313.867	— 400.855	— 38 345.041	— 60 638.515	— 217 677.337

In Tab. XVIII bezeichnen die Zahlen in der 2. Spalte die Werte der auszugleichenden Funktion, in diesem Falle die Sterbenswahrscheinlichkeiten erweitert mit 10^2. Man setzt die Konstante $m = 10.0$ und erhält dadurch die Zahlen in der 3. Spalte, mit welchen dann die folgenden Rechnungen ausgeführt sind. Die unter dem Strich angegebenen Zahlen zeigen die Summen der entsprechenden Spalten. Zwischen diesen muß natürlich folgende Beziehung bestehen:

$$\sum (x+1)^4 (F-m) = \sum x^4 (F-m) + 4 \sum x^3 (F-m) + 6 \sum x^2 (F-m) + 4 \sum x (F-m) + \sum (F-m).$$

Man bekommt

$$\begin{aligned} \sum x^4 (F-m) &= -\ 60\,638.515, \\ 4 \sum x^3 (F-m) &= -\ 153\,380.164, \\ 6 \sum x^2 (F-m) &= -\ 2\,405.180, \\ 4 \sum x (F-m) &= -\ 1\,255.468, \\ \sum (F-m) &= +\ 1.940, \\ \hline \sum &= -\ 217\,677.337 = \sum (x+1)^4 (F-m), \end{aligned}$$

und damit sind alle Rechnungen mit Ausnahme von $\Sigma(F-m)^2$ kontrolliert. Aus den Summen bekommt man nun

$$\nu_2' = \frac{1}{N} \sum (F-m)^2 = 2.0761$$

und weiter nach (44)

$$d_0 = + \ 0.06690,$$
$$d_1 = - \ 10.828,$$
$$d_2 = - \ 13.823,$$
$$d_3 = -1322.25,$$
$$d_4 = -2091.00.$$

Mit Hilfe der oben berechneten α-Konstanten erhält man darauf nach Einsetzung in (43)

$$c_1\beta_1' = - \ 10.823 \text{ sowie hieraus } \beta_1' = - \ 0.1547,$$
$$c_2\beta_2' = - \ 18.506 \quad „ \quad „ \quad \beta_2' = - \ 0.004788,$$
$$c_3\beta_3' = + \ 39.283 \quad „ \quad „ \quad \beta_3' = + \ 0.0001880,$$
$$c_4\beta_4' = + 637.80 \quad „ \quad „ \quad \beta_4' = + \ 0.00005824.$$

Die Dispersion der unausgeglichenen Reihe wird aus

$$\sigma^2 = \nu_2' - d_0^2 = 2.0716$$

berechnet, wonach man die Dispersionen nach Ausgleichung 1., 2., 3. und 4. Grades nach (46) aus resp.

$$\sigma^2_{A_1} = \sigma^2 \ - c_1\beta_1'^2 = 0.3978,$$
$$\sigma^2_{A_2} = \sigma^2_{A_1} - c_2\beta_2'^2 = 0.3096,$$
$$\sigma^2_{A_3} = \sigma^2_{A_2} - c_3\beta_3'^2 = 0.3022,$$
$$\sigma^2_{A_4} = \sigma^2_{A_3} - c_4\beta_4'^2 = 0.2651$$

erhält. Damit ist in einigen Stunden eine numerische Lösung erreicht, die mit der gewöhnlichen Methode eine Arbeit von mehreren Tagen erheischen würde.

Aus obenstehenden σ_A-Werten sind nach (54) die Zahlen der zweiten Spalte in Tab. XIX berechnet worden. In der 3. Spalte bezeichnen die Zahlen entsprechende Q^2_A-Werte für das weibliche Geschlecht und in der 4. die nach (67) berechneten Korrelationskoeffizienten zwischen den Sterbenswahrscheinlichkeiten der Männer und Frauen.

Obwohl die Koeffizienten mit jedem neuen Gliede abnehmen, das in die Gleichung der Regressionslinie eingefügt wird, scheint dennoch die Konvergenz ziemlich gut zu sein, und die säkulare Variation ist praktisch gesehen bereits nach der Ausgleichung 2. Grades eliminiert.

Tab. XIX.

Die Ausgleichung in der Altersstufe 0—1.

Grad der Ausgl.	Q^2_A		Korrel.
	Männer	Frauen	M. × Fr.
0	164.28	143.10	+ 0.991
1	32.71	28.15	+ 0.952
2	26.47	21.66	+ 0.938
3	26.88	20.54	+ 0.937
4	24.57	19.54	+ 0.928

Tab. XX.

Die Ausgleichung in der Altersstufe 4—5.

Grad der Ausgl.	Q^2_A		Korrel.
	Männer	Frauen	M. × Fr.
0	65.49	61.58	+ 0.974
1	9.60	8.54	+ 0.808
2	9.97	8.87	+ 0.808
3	9.89	7.91	+ 0.822
4	10.22	8.24	+ 0.824

In Tab. XX sind entsprechende Zahlen für die Altersstufe 4—5 angegeben. Hier ist die säkulare Abnahme annähernd linear vorsichgegangen. Bei Ausgleichung höheren Grades zeigen die Divergenzkoeffizienten sogar zunehmende Tendenz bei den Männern, ein Zeichen dafür, daß die beste Ausgleichung mit einer Geraden erreicht wird. Wenn die säkulare Variation trotzdem im vorhergehenden konsequent mit Kurven zweiten Grades eliminiert worden ist, so ist das aus praktischen Gründen erfolgt. Die Fehler, die dadurch in den Divergenzkoeffizienten entstehen können, sind aber unbedeutend. Bei $N = 29$ hat man nämlich $Q^2_{A_2} : Q^2_{A_1} = 27\,\sigma^2_{A_2} : 26\,\sigma^2_{A_1}$, und da immer $\sigma^2_{A_2} \leq \sigma^2_{A_1}$, wird $Q^2_{A_2} : Q^2_{A_1} \leq 1.04$, weshalb die Fehler also 4 % niemals übersteigen können. Bei Reihen, die gar keine säkulare Ab- oder Zunahme aufweisen, übersteigen die entsprechenden Fehler nicht 8 %. Derartige Reihen sind aber sehr selten.

In beiden soeben untersuchten Altersstufen waren die Divergenzkoeffizienten so groß, daß über den übernormalen Charakter der Dispersionen kein Zweifel herrschen konnte, selbst wenn die säkulare Variation nicht vollständig eliminiert gewesen wäre. Anders liegen aber die Verhältnisse in den mittleren Altersstufen. Dort haben wir so kleine Q^2_A-Werte, daß es von größter Wichtigkeit ist zu entscheiden, ob nicht die schwache Übernormalität vielleicht auf noch nicht eliminierter säkularer Variation beruhen könne. Ich habe daher ähnliche Untersuchungen für alle Reihen der Altersgruppe 25—30 vorgenommen, der Gruppe, in der die Divergenzkoeffizienten durchschnittlich am kleinsten sind, und wo außerdem die Krümmung der säkularen Regressionslinien am stärksten hervortritt. Die Resultate sind in Tab. XXI A—C zusammengestellt.

Tab. XXI.

Die Ausgleichung in der Gruppe 25—30.

Die Werte von Q^2_A und $\bar{Q}^2_A$.

A.

Männer.

Alter	Grad der Ausgleichung				
	0	1	2	3	4
25—26	1.81	1.62	1.64	1.68	1.60
26—27	1.60	1.57	1.29	1.32	1.20
27—28	2.06	1.51	1.13	1.14	1.19
28—29	1.73	1.18	1.23	1.13	1.12
29—30	1.90	1.81	1.76	1.79	1.79
$\bar{Q}^2_A$	1.82	1.54	1.41	1.41	1.38

B.

Alter	Grad der Ausgleichung				
	0	1	2	3	4
25—26	1.27	1.28	1.30	1.33	1.27
26—27	1.61	1.56	0.90	0.94	0.96
27—28	1.28	1.16	1.21	1.25	1.18
28—29	1.55	1.51	1.40	1.43	1.48
29—30	1.33	0.94	0.90	0.94	0.89
$\bar{Q}^2_A$	1.41	1.29	1.14	1.18	1.16

C.

Die Werte von $\bar{r}_A$ und γ_A.

Grad der Ausgleichung	$\bar{r}_A$		γ_A
	Männer	Frauen	
0	+ 0.429	+ 0.270	+ 0.373
1	+ 0.332	+ 0.198	+ 0.288
2	+ 0.292	+ 0.139	+ 0.224
3	+ 0.310	+ 0.134	+ 0.232
4	+ 0.276	+ 0.143	+ 0.222

Aus den Tabellen A und B ist ersichtlich, daß die säkulare Variation in fast allen Reihen bei Männern wie bei Frauen als vollständig eliminiert betrachtet werden kann, wenn die Ausgleichung bis zum 2. Grade durchgeführt wird. Der durchschnittlich übernormale Charakter der Dispersionen steht danach fest. In Tab. C sind die $\bar{r}_A$-Werte für Männer und Frauen, nach (90) berechnet, wiedergegeben, sowie die nach (98) bestimmten γ_A-Koeffizienten. Auch diese Zahlen scheinen nach vorgenommener Ausgleichung 2. Grades vom Einfluß der säkularen Variation befreit zu sein.

In den übrigen Gruppen der mittleren Altersstufen schließen sich die Regressionslinien 2. Grades mindestens gleich gut den beobachteten Sterbenswahrscheinlichkeiten an, wie in dieser Gruppe, und ich habe es deshalb nicht für nötig befunden, weitere Untersuchungen für diese Altersstufen durchzuführen. In der männlichen Altersgruppe 55—60 sind dagegen die Sterblichkeitskurven so unregelmäßig, daß es schwierig ist, die Zuverlässigkeit der Ausgleichung nur mit Augenmaß zu entscheiden. Aus diesem Grunde habe ich sie auch für diese Gruppe numerisch untersucht. Die Ergebnisse sind in Tab. XXII A und B wiedergegeben.

Tab. XXII.

Die Ausgleichung in der männlichen Gruppe 55—60.

A.

Die Werte von Q^2_A und $\bar{Q}^2_A$.

Alter	Grad der Ausgleichung				
	0	1	2	3	4
55—56	2.84	2.31	2.40	2.50	2.40
56—57	2.65	1.86	1.91	1.70	1.76
57—58	4.00	2.89	2.64	2.48	2.25
58—59	1.83	1.15	1.18	1.10	1.13
59—60	2.95	2.24	2.19	2.27	2.29
$\bar{Q}^2_A$	2.86	2.09	2.06	2.01	1.97

B.

Die Werte von $\bar{r}_A$.

Grad der Ausgleichung				
0	1	2	3	4
+ 0.533	+ 0.340	+ 0.370	+ 0.359	+ 0.349

Auf die Divergenzkoeffizienten der einzelnen Reihen übt die säkulare Variation einen ziemlich verschiedenartigen Einfluß aus, doch zeigen die Ziffern im Durchschnitt eine starke Konvergenz. Die säkulare Abnahme kann daher auch hier in völlig befriedigender Weise vermittels einer Parabel 2. Grades eliminiert werden.

Schließlich ist die Effektivität der Ausgleichung auch in der Gruppe 75—80 untersucht worden. Zwar scheint die säkulare Abnahme hier fast linear vonstatten gegangen zu sein; die Sterblichkeitskurven enthalten aber sämtlich einige ziemlich deutlich ausgeprägte Maxima, die bei Ausgleichung höheren Grades möglicherweise auf den Gang der Regressionslinie einwirken könnten. Daß dies jedoch nicht der Fall ist, geht aus den Zahlen in Tab. XXIII A—C hervor. Sowohl die Divergenz- als auch die Korrelationskoeffizienten zeigen nämlich, daß die möglichst beste Ausgleichung durch eine Gerade erreicht wird. Der Umstand jedoch, daß die Elimination im vorhergehenden mit einer Kurve 2. Grades vorgenommen wurde, hat auf die Gültigkeit der Ergebnisse nicht den geringsten Einfluß.

Tab. XXIII.

Die Ausgleichung in der Gruppe 75—80.

Die Werte von Q^2_A und $\overline{Q}^2_A$.

A.

Männer.

Alter	Grad der Ausgleichung				
	0	1	2	3	4
75—76	4.04	2.06	2.13	2.21	1.83
76—77	5.12	3.08	3.18	3.14	3.10
77—78	4.37	2.90	2.99	2.96	2.96
78—79	3.60	2.22	2.16	2.09	2.05
79—80	3.45	2.78	2.88	2.71	2.77
$\overline{Q}^2_A$	4.12	2.61	2.67	2.62	2.54

B.

Alter	Grad der Ausgleichung				
	0	1	2	3	4
75—76	6.46	4.89	5.03	5.13	5.14
76—77	4.63	3.39	3.52	3.47	3.59
77—78	4.91	3.68	3.82	3.97	4.13
78—79	5.69	3.85	3.95	3.97	4.04
79—80	5.16	4.29	4.12	4.29	4.22
$\overline{Q}^2_A$	5.37	4.02	4.09	4.17	4.22

C.

Die Werte von $\overline{r}_A$ und r_A.

Grad der Ausgleichung	$\overline{r}_A$		r_A
	Männer	Frauen	
0	+ 0.676	+ 0.771	+ 0.712
1	+ 0.503	+ 0.690	+ 0.591
2	+ 0.512	+ 0.710	+ 0.603
3	+ 0.489	+ 0.714	+ 0.602
4	+ 0.501	+ 0.729	+ 0.628

In allen hier untersuchten Altersstufen kann demnach die säkulare Variation für die Zeit von 1886—1914 in völlig befriedigender Weise durch eine Parabel 2. Grades eliminiert werden, und dieses dürfte mit großer Bestimmtheit auch für die übrigen Altersstufen zutreffen. Bei den beiden Teilperioden schließen sich die linearen Regressionslinien im allgemeinen der parabolischen Kurve sehr wohl an, und, obwohl diese Perioden zu kurz sind, um eine Ausgleichung höheren Grades zu erlauben, kann man wohl dennoch annehmen, daß die säkulare Variation auch hier befriedigend eliminiert ist. Die Ergebnisse der vorangehenden Kapitel gelten also in demselben Maße, kann man sagen, wie wenn keine säkulare Ab- oder Zunahme stattgefunden hätte.

Ich kann in diesem Zusammenhange nicht umhin, einige von den Resultaten mitzuteilen, die ich bei Versuchen, die säkulare Variation nach der „variate-difference"-Methode zu eliminieren, erhalten habe. Diese sind nämlich sehr unbefriedigend ausgefallen und lassen die Brauchbarkeit der Methode äußerst zweifelhaft erscheinen.

In Tab. XXIV A zeigen sowohl die Divergenz- als auch die Korrelationskoeffizienten eine fortwährende Zunahme[1]). Man könnte danach zu der Ansicht gelangen, daß die säkulare Variation noch bei der 6. Differenz, entsprechend einer Ausgleichung mit einer Parabel 5. Grades, nicht eliminiert sei, während die Zahlen in Tab. XIX angeben,

[1]) Dispersionen und Korrelationskoeffizienten sind nach den von ANDERSON in „Biometrika", Vol. X, pag. 269, angegebenen Formeln berechnet.

Tab. XXIV.

Ausgleichung nach der „variate-difference"-Methode.

A.

Die Altersstufe 0—1.

Differenz	Q^2		Korrel. M. × Fr.
	Männer	Frauen	
0	164.28	143.10	+ 0.991
1	26.75	24.57	+ 0.945
2	28.66	27.24	+ 0.957
3	31.34	29.91	+ 0.963
4	34.25	32.42	+ 0.967
5	37.07	34.55	+ 0.972
6	39.81	36.37	+ 0.975

B.

Die Altersstufe 4—5.

Differenz	Q^2		Korrel. M. × Fr.
	Männer	Frauen	
0	65.49	61.58	+ 0.974
1	6.99	6.10	+ 0.693
2	6.17	5.27	+ 0.504
3	6.14	4.86	+ 0.406
4	6.23	4.51	+ 0.324
5	6.51	4.30	+ 0.268
6	6.89	4.20	+ 0.332

daß dieses praktisch schon nach der Ausgleichung 2. Grades der Fall ist. Aber, auch wenn die vorangehenden Ergebnisse nicht bekannt wären, würde der Gang der Zahlen schon genügen, um über die Gültigkeit der Methode Zweifel aufkommen zu lassen. Nach den Voraussetzungen, auf welchen sie fußt, ist es nämlich theoretisch undenkbar, daß die Divergenzkoeffizienten, wenn sie im Steigen begriffen sind, konstanten Werten entgegengehen können. Das kann folgendermaßen bewiesen werden. $x_1, x_2, \ldots x_N$ sei die Reihe, die ausgeglichen werden soll, und $X_1, X_2, \ldots X_N$ die zeitfreien Residualen[1]). Nehmen wir nun an, daß die säkulare Variation nach der r^{ten} Differenz vollständig eliminiert sei, dann ist

$$\Delta_r x_k = \Delta_r X_k, \tag{125}$$

während für Differenzen niedriger Ordnung

$$\Delta_m x_k = \Phi_m(k) + \Delta_m X_k, \qquad (m < r) \tag{126}$$

wo $\Phi_m(k)$ eine ganze rationale Funktion von k ist.

Unter der Bedingung, daß

$$M(X_k) = 0,$$

$$M(X_k X_l) = \begin{cases} 0, & \text{wenn } k \neq l, \\ \sigma^2_X & \text{„ } k = l, \end{cases}$$

hat ANDERSON[2]) nachgewiesen, daß

$$M(\Delta_r X_k) = 0,$$

$$M(\Delta^2_r X_k) = \frac{2\,r!}{r!\,r!}\,\sigma^2_X.$$

[1]) Den Ausdruck „die zeitfreien Residualen" fasse ich hier etwas allgemeiner als vorher.

[2]) „Biometrika", loc. cit.

Bezeichnet man die Dispersion von $\Delta_r x_k$ mit σ_{Δ_r}, dann bekommt man aus (125) und (126)

$$\frac{r!\,r!}{2\,r!}\,M(\sigma^2_{\Delta_r}) = \sigma^2 x,$$

$$\frac{m!\,m!}{2\,m!}\,M(\sigma^2_{\Delta_m}) = \sigma^2 x + c^2 m$$

und demnach

$$\frac{r!\,r!}{2\,r!}\,M(\sigma^2_{\Delta_r}) \leq \frac{m!\,m!}{2\,m!}\,M(\sigma^2_{\Delta_m}).$$

Da nun die Divergenzkoeffizienten hier durch

$$Q^2_{A_r} = \frac{r!\,r!}{2\,r!}\,\sigma^2_{\Delta_r} : w_2$$

definiert sind, so muß auch die Ungleichung

$$Q^2_{A_r} \leq Q^2_{A_m}$$

innerhalb der Grenzen der zufälligen Fehler gelten.

Selbst wenn die Koeffizienten in Tab. XXIII A für die folgenden Differenzen einigermaßen konstante Werte bekommen sollten, kann also wegen des Zahlenverlaufs behauptet werden, daß die „variate-difference"-Methode hier für die Elimination der säkularen Variation nicht angewandt werden kann. Der Grund dafür ist wohl darin zu suchen, daß die zeitfreien Residualen nicht, wie die Methode voraussetzt, voneinander unabhängig sind[1]). Nun setzt freilich auch die Methode der kleinsten Quadrate die Unabhängigkeit der zeitfreien Residualen voneinander voraus; aber diese Bedingung braucht hier nur im Durchschnitt erfüllt zu sein, und auch wenn dieses nicht der Fall wäre, bestimmt die Methode der kleinsten Quadrate immer die Kurve, die die Quadratsumme der Abweichungen zu einem Minimum macht. Es erscheint dagegen sehr schwer, bei entsprechenden Verhältnissen in der „variate-difference"-Methode einen greifbaren Haltpunkt zu finden.

In Tab. XXIV B verraten die Koeffizienten im Gegensatz zu den früheren eine abnehmende Tendenz, scheinen aber noch bei der 6. Differenz nicht konstante Werte erreicht zu haben, und zwar, obwohl die säkulare Abnahme hier, entsprechend den Zahlen in Tab. XX, beinahe linear war. Auch in der Altersgruppe 25—30 sind ein paar Reihen mit ähnlichen Ergebnissen untersucht worden. Die Methode scheint daher für die Elimination der säkularen Variation in der Sterblichkeit vollkommen unbrauchbar zu sein.

Es würde uns über den Rahmen dieser Arbeit hinausführen, den äußeren Faktoren nachzugehen, die auf die Sterblichkeit einwirken und die die Fluktuationen der Grundwahrscheinlichkeiten um die säkularen Regressionslinien veranlassen, und solche

[1]) BROWNLEE hat gerade in diesem Punkt die Methode einer scharfen Kritik unterworfen. Journ. of the R. Stat. Soc. Vol. 80, pag. 222.

Versuche sollen auch nicht unternommen werden. Über einen Punkt können jedoch aus „Sveriges Officiella Statistik“ direkte, wenn auch unvollständige Angaben erhalten werden, nämlich über den Einfluß der sog. Seuchensterblichkeit auf die totale Sterblichkeit. Da es von gewissem Interesse sein kann zu erfahren, in welchem Grade die Stabilitäts- und Korrelationsmaße dadurch beeinträchtigt werden, möchte ich folgende kleine Untersuchung wiedergeben.

In den offiziellen Veröffentlichungen wird die Seuchensterblichkeit mit Hilfe der Sterbeziffer für die Altersgruppen 0—10 und 10—ω angegeben. Die Untersuchungen sind nur für die Gruppe 0—10 vorgenommen und auf die Zeit von 1886—1900 beschränkt worden[1]). Zu den Seuchen werden in dieser Altersgruppe hauptsächlich die epidemischen Kinderkrankheiten gezählt.

Bezeichnet man

die totale Sterblichkeit mit ..... p,
„ Seuchensterblichkeit „ ..... p_1,
„ Seuchensterblichkeit—totale Sterblichkeit mit ..... p_2,

so ist

$$p_2 = p - p_1.$$

Die säkulare Variation wird linear angenommen, und für p, p_1 und p_2 ergeben sich folgende Charakteristiken:

Tab. XXV.

Charakteristiken	Männer			Frauen		
	p	p_1	p_2	p	p_1	p_2
$10^3 p'$	24.28	6.42	17.86	21.40	6.25	15.15
$10^2 v$	— 1.52	— 3.16	— 0.93	— 1.50	— 3.25	— 0.92
Q^2_A	66.79	74.82	23.35	54.41	65.91	17.15
$100 \rho_A$	6.87	14.29	4.69	6.96	13.77	4.39
r_A		+ 0.614			+ 0.655	

Der Durchschnittswert der Seuchensterblichkeit beläuft sich also bei Männern auf fast 25 % des Durchschnittswertes der totalen Sterblichkeit, bei Frauen auf etwas mehr. Die Seuchensterblichkeit p_1 zeigt aber eine bedeutend schnellere säkulare Abnahme, die, mit dem relativen Maß $10^2 v$ gemessen, doppelt so stark ist, wie die Abnahme von p und mehr als dreimal so stark, wie die von p_2. Auf die Stabilität übt die Seuchensterblichkeit einen erheblichen Einfluß aus, und sowohl Divergenz- als auch Störungskoeffizienten werden viel kleiner für p_2 als für p. Die stark übernormalen Dispersionen in den Kinderjahren können demnach zum großen Teil dem störenden Einfluß der Seuchensterblichkeit zugeschrieben werden. Der Koeffizient r_A, nach (67) bestimmt, deutet eine ziemlich starke positive Korrelation zwischen p_1 und p_2 an, d. h. zwischen der Seuchensterblichkeit einerseits und der totalen Sterblichkeit minus der Seuchen-

[1]) Nach 1910 wird die Seuchensterblichkeit nicht mehr genannt, weshalb auch die Untersuchung für den ganzen Zeitraum 86—14 nicht durchgeführt werden kann.

sterblichkeit andererseits. Hätten p_1 und p_2 beide normale Dispersionen, so wäre die Korrelation theoretisch durch

$$r = -\sqrt{\frac{p_1 p_2}{(1-p_1)(1-p_2)}}$$

gegeben und würde somit negativ ausfallen. Die stark positiven Werte von den Koeffizienten können indessen z. T. dadurch erklärt werden, daß der Sterblichkeitsquotient p_2 nicht vollständig vom Einfluß der Seuchensterblichkeit befreit worden ist.

Die Korrelationen zwischen entsprechenden männlichen und weiblichen Sterblichkeitsquotienten werden durch folgende Zahlen ausgedrückt:

für p ..... $r_A = + 0.970$,
„ p_1 ..... $= + 0.989$,
„ p_2 ..... $= + 0.924$.

Die Korrelation wird also für p_2 etwas schwächer als für p, ist aber nicht in demselben Grade von der Seuchensterblichkeit beeinflußt, wie die Stabilität.

Es muß aber noch einmal betont werden, daß die offiziellen Angaben für eine sichere Beurteilung dieser Fragen nicht hinreichend genau sind.

Inhaltsverzeichnis.

		Seite
Einleitung		3
Kap. I.	Die LEXISsche Reihe und die Stabilitätsmaße	5
„ II.	Die Stabilitätsmaße nach Elimination der säkularen Variation	10
„ III.	Die Korrelation zwischen zwei von der säkularen Variation befreiten Reihen	22
„ IV.	Die Stabilitätsuntersuchung	27
„ V.	Die Korrelationsuntersuchung	40
„ VI.	Die Effektivität der Ausgleichung. Die Einwirkung der Seuchensterblichkeit auf Stabilität und Korrelation	60
„ VII.	Zusammenfassung der numerischen Resultate	70

DRUCK VON LÜTCKE & WULFF, HAMBURG